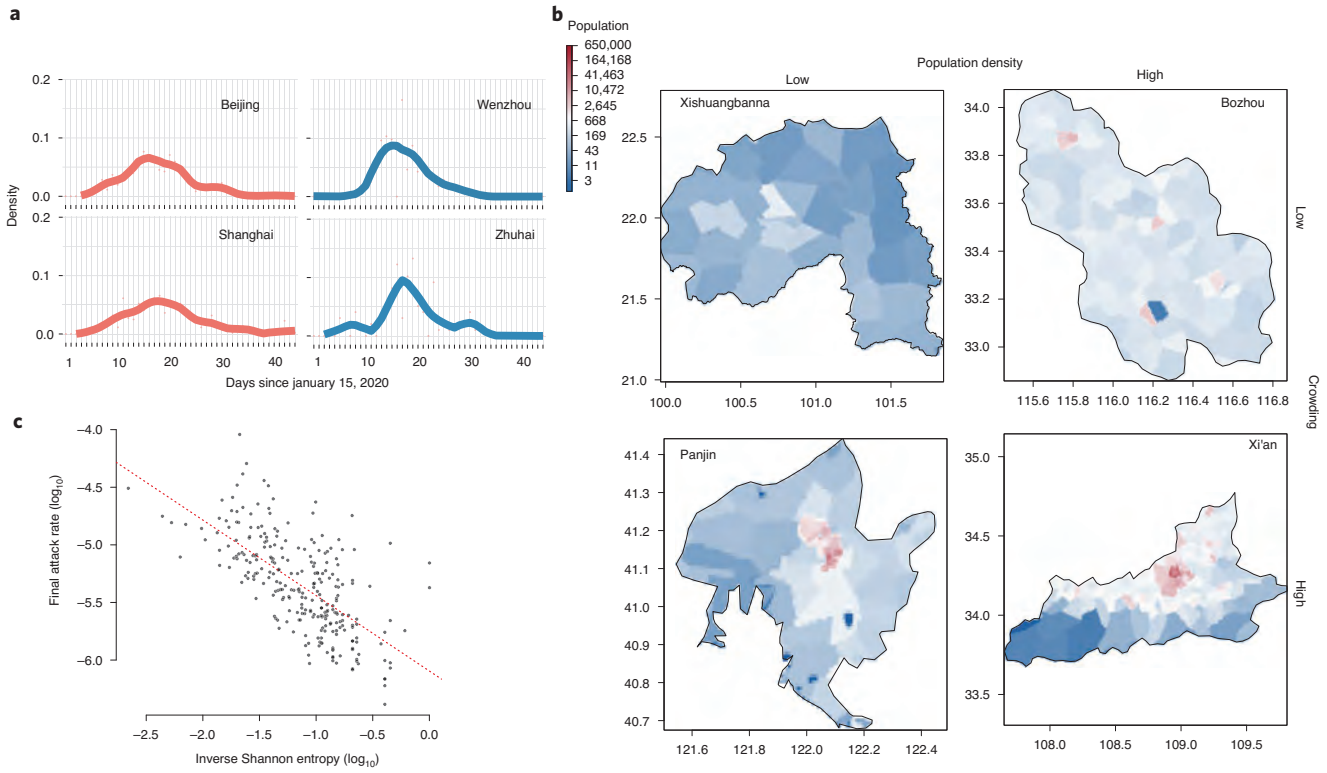


Fig. 1 | Maps of crowding in prefectures in China. a, Examples of epidemic curves that are normalized to show the percentage of cases across the whole epidemic that occur at each given day. Beijing and Shanghai (red) have less peaked epidemics than Wenzhou and Zhuhai. **b**, Examples of prefectures in China with different levels of crowding and population size. The color scale illustrates the estimated number of inhabitants per grid cell (1 km × 1 km). **c**, Relationship between the Shannon index of the incidence curve and the final attack rate for prefectures in China.

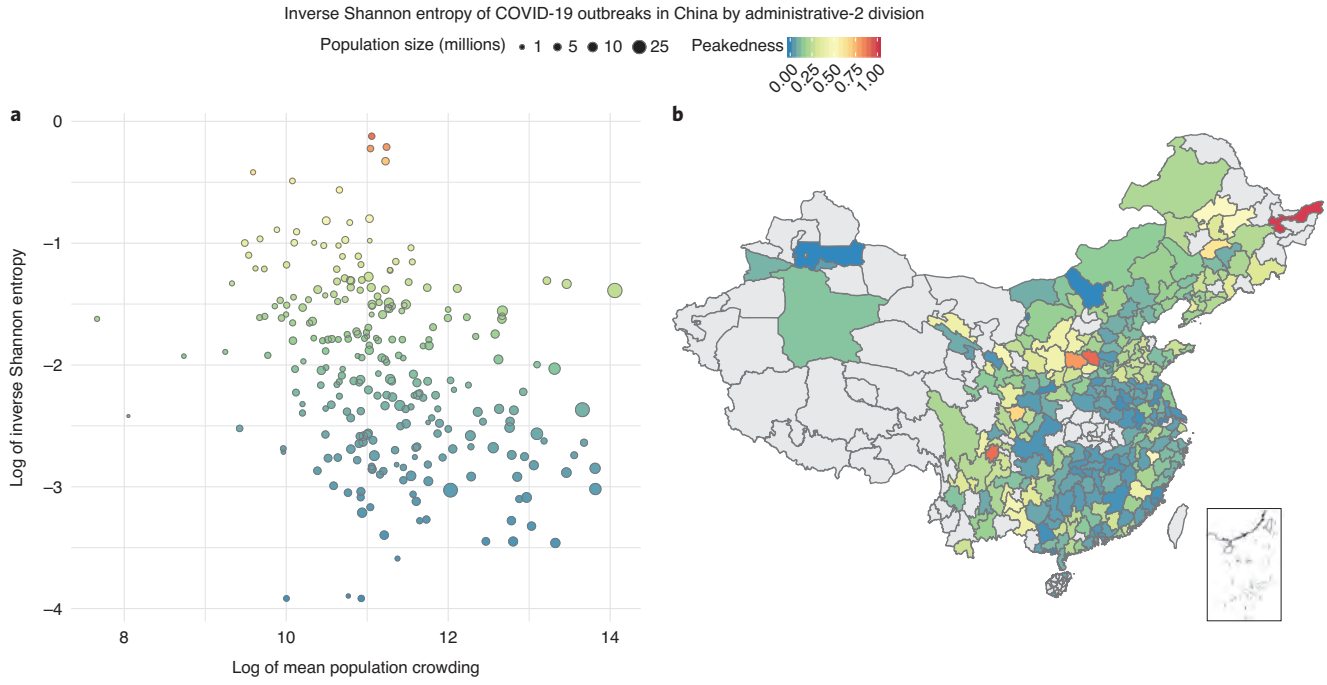


Fig. 2 | Crowding and the temporal clustering of transmission of COVID-19 in China. a, Negative association between $\log_{10}$ of epidemic peakedness, as measured by Shannon's diversity index, and log population crowding, as measured by Lloyd's mean crowding. The point sizes indicate the size of the population in each city. **b**, Map of epidemic peakedness in China at the prefectural level. Blue and green colors indicate lower peakedness; red and yellow colors indicate higher peakedness. Gray prefectures had either no reported cases or were not included in analyses due to potential inconsistencies in reporting of early cases (Hubei Province).

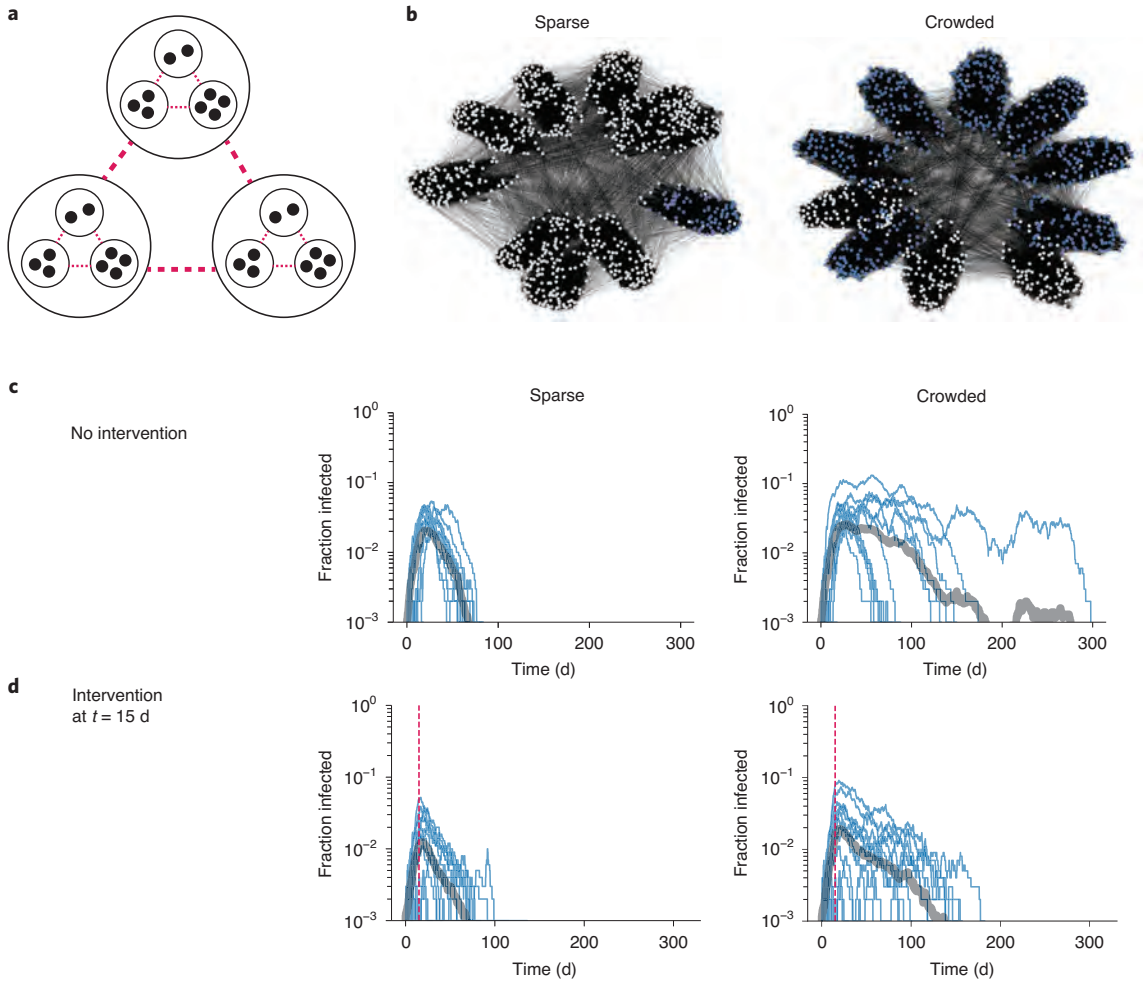


Fig. 3 | Mechanisms generating less peaked epidemics in crowded populations. a, Schematic of a hierarchically structured population model consisting of households and 'neighborhoods' within a prefecture. Transmission is more likely among contacts connected at lower spatial levels. Crowded populations have a greater number of contacts outside the household, and interventions reduce the number of these connections in both populations (pink dotted lines). **b**, **c**, Simulated outbreak dynamics in the absence of interventions in crowded versus sparse populations. For the networks in **b**, blue nodes are individuals who were eventually infected by the end of the outbreak. In **c**, thin blue lines show individual realizations of the model, the average shown by the thick gray line. **d**, Simulated outbreak dynamics in the presence of strong social distancing measures in crowded versus sparse populations. The intervention was implemented at day 15 (vertical dotted line) and led to a 75% reduction in contacts, similar to observed changes in contact rates in China[35,36]. Mean values of median log epidemic peakedness (Shannon index) are −2.3 for low crowding and −2.8 for high crowding.

Simulation of COVID-19 epidemics in hierarchically structured populations. We hypothesize that the mechanism underlying our central observation—that more crowded cities experience less peaked outbreaks—is that crowding enables sustained transmission among households and through a city's population, leading incidence to be widely distributed through time. To explore this proposed mechanism, we simulated stochastic epidemic dynamics in two types of populations. Simple, well-mixed transmission models in which contact rates are high in crowded regions were not consistent with our findings, because they predict that crowded regions would have more temporally clustered outbreaks. To capture realistic contact patterns, we created hierarchically structured populations[29] in which individuals had high rates of contact within their social units (which are defined broadly and could represent households, care homes, hospitals, prisons, etc); lower rates with individuals from other units but within the same neighborhoods; and relatively rare contact with other individuals in other neighborhoods within the same prefecture (Fig. 3a). These assumptions are consistent with reports that most onward transmission after lockdowns were implemented occurred in households or in other close-contact situations[2,30]. In this scenario, less crowded prefectures often had more peaked and shorter outbreaks that were isolated to specific neighborhoods, whereas more crowded prefectures could sustain drawn-out outbreaks of larger final size, which jumped among the more highly connected neighborhoods (Fig. 3b,c). Further, if the reproduction number of COVID-19 is over-dispersed[31–33], then crowding could enable local outbreaks to spread more widely due to the availability of contacts[34].

We also simulated outbreak dynamics under extensive social distancing measures, as observed in Chinese prefectures (75% reduction in contact rates[35,36]). If social distancing reduces non-household contacts by the same relative amount in all locations, there will be more contacts remaining in crowded areas, because baseline contact rates are higher. Consequently, outbreaks in crowded regions could be larger and take longer to end after intervention (Fig. 3d, Fig. 1c and Extended Data Fig. 1).

Using the fitted model from China paired with globally comprehensive covariates, we extrapolated our results to cities across the world (Fig. 4). Human mobility data from Baidu were not available for locations outside of China. Therefore, we used aggregated

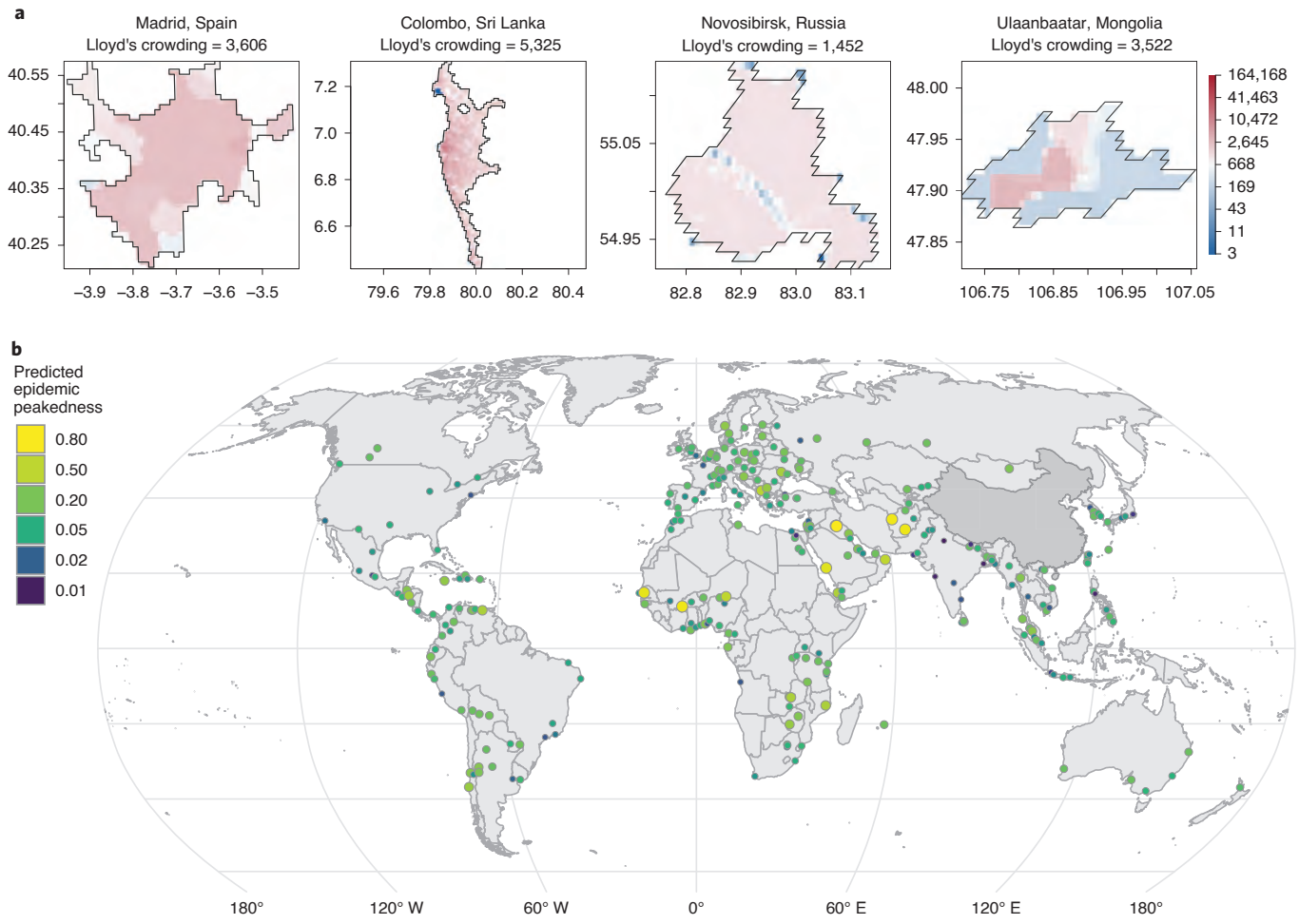


Fig. 4 | Predicted epidemic peakedness across the world. a, Maps of cities and their population densities at a 1 × 1-km scale. Madrid, Spain, and Colombo, Sri Lanka, have low predicted peakedness, whereas Novosibirsk, Russia, and Ulaanbaatar, Mongolia, have high predicted peakedness. **b**, Map of predicted epidemic peakedness for 310 cities across the world for which both human population data and mobility data were available for the study period.

human mobility data from Google's COVID Mobility Research Dataset (Methods) to capture relative differences in human mobility through time. At the global scale, cities in yellow are predicted to have concentrated and peaked epidemics, whereas cities in blue are predicted to have more prolonged outbreaks (Fig. 4b; a full list is provided in the Supplementary Information). In general, the epidemics in coastal cities were less peaked and were larger and more prolonged, which could be attributable to high levels of population crowding in coastal cities. These predictions rely on fitted relationships of the first epidemic curves from Chinese and Italian cities and, therefore, should be interpreted very cautiously when generalizing to other settings.

Discussion

Our findings confirm previous work on the peakedness of epidemics transmission for influenza in cities[13]. Our work provides empirical support for the role of spatial organization in determining infectious disease dynamics[29,37] and, specifically, spatial variability in transmission parameters[38]. Furthermore, with lower total incidence in small cities compared to larger cities, the risk of resurgence could be elevated owing to lower population immunity after the first wave of the epidemic. Higher seroprevalence for COVID-19 in urban areas[39] provides initial data to support these findings; however, there remains an urgent need to expand serological data collection and provide a full picture of attack rates across cities worldwide[40]. Even though our model does not account for over-dispersion in COVID-19 transmission, there is a theoretical link between the reproduction number in heterogeneous environments and Lloyd's crowding index of aggregation[41], such that the reproduction number increases with higher aggregation[34]. We report that, in dense cities, reductions in mobility tend to be larger, which potentially elevates the effectiveness of non-pharmaceutical interventions in dense cities[42]. However, assessing the effect of within-city connectivity and its spatial heterogeneity on disease dynamics will be critical to further our understanding of how COVID-19 spreads in urban areas. We found that there is an association between climatic factors and the peakedness of epidemics, but particular caution will need to be applied in interpreting these relationships outside the two studied countries (Italy and China). More work is needed to provide causal evidence for the effect of climatic factors on transmission dynamics of COVID-19 during the pandemic and post-pandemic phases[10].

Currently, non-pharmaceutical interventions are the primary control strategy for COVID-19. As a result, public health measures are often focused on 'flattening the curve' to lower the risk of essential services running out of capacity. We show that spatial context, especially crowding, are important factors for assessing the shape of epidemic curves. Therefore, it will be critical to view non-pharmaceutical interventions through the perspective of

crowding—that is, how does an intervention reduce the circle of contacts of an average individual—in cities across the world.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-020-1104-0.

Received: 29 June 2020; Accepted: 16 September 2020;
Published online: 5 October 2020

References

1. Fraher, E. P. et al. Ensuring and sustaining a pandemic workforce. *N. Engl. J. Med.* **382**, 2181–2183 (2020).
2. Leung, K., Wu, J. T., Liu, D. & Leung, G. M. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. *Lancet* **395**, 1382–1393 (2020).
3. Ji, Y., Ma, Z., Peppelenbosch, M. P. & Pan, Q. Potential association between COVID-19 mortality and health-care resource availability. *Lancet Glob. Health* **8**, e480 (2020).
4. Rosenbaum, L. Facing Covid-19 in Italy—ethics, logistics, and therapeutics on the epidemic's front line. *N. Engl. J. Med.* **382**, 1873–1875 (2020).
5. Tian, H. et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* **368**, 638–642 (2020).
6. Kraemer, M. U. G. et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497 (2020).
7. Lipsitch, M., Swerdlow, D. L. & Finelli, L. Defining the epidemiology of Covid-19—studies needed. *N. Engl. J. Med.* **382**, 1194–1196 (2020).
8. World Health Organization. Coronavirus disease 2019 (COVID-19) Situation Report - 71 https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200331-sitrep-71-covid-19.pdf?sfvrsn=4360e92b_8 (2020).
9. Zhao, S. et al. Quantifying the association between domestic travel and the exportation of novel coronavirus (2019-nCoV) cases from Wuhan, China in 2020: a correlational analysis. *J. Travel Med.* **27**, 1–3 (2020).
10. Baker, R. E., Yang, W., Vecchi, G. A., Metcalf, C. J. E. & Grenfell, B. T. Susceptible supply limits the role of climate in the COVID-19 pandemic. *Science* **369**, 315–319 (2020).
11. Rocklöv, J. & Sjödin, H. High population densities catalyse the spread of COVID-19. *J. Travel Med.* **27**, taaa038 (2020).
12. Kraemer, M. U. G. et al. Big city, small world: density, contact rates, and transmission of dengue across Pakistan. *J. R. Soc. Interface* **12**, 20150468 (2015).
13. Dalziel, B. D. et al. Urbanization and humidity shape the intensity of influenza epidemics in U.S. cities. *Science* **362**, 75–79 (2018).
14. Shaman, J., Pitzer, V. E., Viboud, C., Grenfell, B. T. & Lipsitch, M. Absolute humidity and the seasonal onset of influenza in the continental United States. *PLoS Biol.* **8**, e1000316 (2010).
15. Gog, J. R. et al. Spatial transmission of 2009 pandemic influenza in the US. *PLoS Comput. Biol.* **10**, e1003635 (2014).
16. Shaman, J. & Kohn, M. Absolute humidity modulates influenza survival, transmission, and seasonality. *Proc. Natl Acad. Sci. USA* **106**, 3243–3248 (2009).
17. Chetty, R. et al. The association between income and life expectancy in the United States, 2001–2014. *JAMA* **315**, 1750–1766 (2016).
18. Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H. & Lipsitch, M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **21**, 1–9 (2020).
19. Crawford, J. M. et al. Laboratory surge response to pandemic (H1N1) 2009 outbreak, New York City Metropolitan Area, USA. *Emerg. Infect. Dis.* **16**, 8–13 (2010).
20. Grasselli, G., Pesenti, A. & Cecconi, M. Critical care utilization for the COVID-19 outbreak in Lombardy, Italy. *JAMA* **323**, 1545–1546 (2020).
21. Li, Q. et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020).
22. Xu, B. et al. Epidemiological data from the COVID-19 outbreak, real-time case information. *Sci. Data* **7**, 106 (2020).
23. Xu, B. et al. Epidemiological data from the COVID-19 outbreak, real-time case information. *figshare* https://doi.org/10.6084/m9.figshare.11949279 (2020).
24. Xu, B. & Kraemer, M. U. G. Open access epidemiological data from the COVID-19. *Lancet Infect. Dis.* **3099**, 30119 (2020).
25. Aurora Big Data. *2017 Mobile Map App Research Report: Which of the Highest, the Baidu, and Tencent Is Strong?* https://baijiahao.baidu.com/s?id=1590386747028939917&wfr=spider&for=pc. (2017)
26. Lloyd, M. 'Mean crowding'. *J. Anim. Ecol.* **36**, 1–30 (1967).
27. May, R. M. & Anderson, R. M. Spatial heterogeneity and the design of immunization programs. *Math. Biosci.* **72**, 83–111 (1984).
28. Anderson, R. M. & May, R. M. *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, 1991).
29. Watts, D. J., Muhamad, R., Medina, D. C. & Dodds, P. S. Multiscale, resurgent epidemics in a hierarchical metapopulation model. *Proc. Natl Acad. Sci. USA* **102**, 11157–11162 (2005).
30. Report of the WHO–China Joint Mission on Coronavirus Disease 2019 (COVID-19) https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf 16–24 (2020).
31. Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. & Getz, W. M. Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359 (2005).
32. Kucharski, A. J. et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect. Dis.* **3099**, 1–7 (2020).
33. Riou, J. & Althaus, C. L. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Euro. Surveill.* **25**, 1–5 (2020).
34. Southwood, T. R. in *Ecological Methods* (ed Southwood, T. R.) 7–69 (Springer Netherlands, 1978).
35. Zhang, J. et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* **368**, 1481–1486 (2020).
36. Lai, S. et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature* **585**, 410–413 (2020).
37. Sattenspiel, L. Simulating the effect of quarantine on the spread of the 1918–19 flu in central Canada. *Bull. Math. Biol.* **65**, 1–26 (2003).
38. Meyers, L. A. Contact network epidemiology: bond percolation applied to infectious disease prediction and control. *Bull. Am. Math. Soc.* **44**, 63–87 (2006).
39. Kissler, S. M. et al. Reductions in commuting mobility predict geographic differences in SARS-CoV-2 prevalence in New York City. *Nat. Commun.* **16**, 4674 (2020).
40. Lipsitch, M., Swerdlow, D. L. & Finelli, L. Defining the epidemiology of Covid-19 — studies needed. *N. Engl. J. Med.* **382**, 1194–1196 (2020).
41. Mat, N. F. C., Edinur, H. A., Razab, M. K. A. A. & Safuan, S. A single mass gathering resulted in massive transmission of COVID-19 infections in Malaysia with further international spread. *J. Travel Med.* **27**, taaa059 (2020).
42. Flaxman, S. et al. Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries. *Nature* **584**, 257–261 (2020).

Methods

Epidemiological data. No officially reported line list was available for cases in China. We used a standardized protocol[43] to extract individual-level data from December 1, 2019, to March 30, 2020. Sources were mainly official reports from provincial, municipal or national health governments. Data included basic demographics (age and sex), travel histories and key dates (dates of onset of symptoms, hospitalization and confirmation). Data were entered by a team of data curators on a rolling basis, and technical validation and geo-positioning protocols were applied continuously to ensure validity. A detailed description of the methodology is available[22]. Lastly, total numbers were matched with officially reported data from China and other government reports. Daily case counts from Italian provinces ($n = 107$) were extracted from the Presidenza del Consiglio dei Ministri Dipartimento della Protezione Civile (https://github.com/pcm-dpc/COVID-19).

Estimating epidemic peakedness. Epidemic peakedness was estimated for each prefecture by calculating the inverse Shannon entropy of the distribution of COVID-19 cases. Inverse Shannon entropy was used to fit time series of other respiratory infections (influenza)[13]. The inverse Shannon entropy of incidence for a given prefecture in 2020 is then given by $v_j = \left(-\sum_i p_{ij} \log p_{ij}\right)^{-1}$. Because v_j is a function of incidence distribution in each location rather than raw incidence, it is invariant under differences in overall reporting rates between cities or attack rates. We then assessed how peakedness $v \propto \sum_j v_j$ varied across geographic areas in China. As an alternative measure of temporal clustering of cases, we estimated the proportion of cases at the peak ± 1 d (Extended Data Fig. 2).

Proxies for COVID-19 interventions using within-city human mobility data from China. Estimates of within-city reductions of human mobility between the period before and after the lockdown was implemented on January 23, 2020, were extracted from Lai et al.[36]. Daily measures of human mobility were extracted from the Baidu Qianxi web platform to estimate the proportion of daily movement within prefectures in China. Relative mobility volume was available from January 2, 2020, to January 25, 2020. For each city, change in relative mobility was defined by $m_i = m_{il}(lockdown)/m_{ib}(baseline)$, where m_i is defined as mobility in prefecture i. Baidu's mapping service is estimated to have a 30% market share in China, and more data can be found[5,6].

Data on drivers of transmission of COVID-19. Prefecture-specific population counts and densities were derived from the 2020 Gridded Population of The World, a modeled continuous surface of population estimated from national census data and the United Nations World Population Prospectus[44]. Population counts are defined at a 30-arc-second resolution (approximately 1 km × 1 km at the equator) and extracted within administrative 2 level cartographic boundaries defined by the National Bureau of Statistics of China. Lloyd's mean crowding, $\frac{\left[\sum_i (q_i - 1) q_i\right]}{\sum_i q_i}$, was estimated for each prefecture, where q_i represents the population count of each non-zero pixel within a prefecture's boundary and the resulting value estimates an individual's mean number of expected neighbors[13,45]. When fitting the models, we consider the numerator $\left[\sum_i (q_i - 1) q_i\right]$, which we refer to as 'contacts', and the denominator $\sum_i q_i$ (that is, population size) as separate predictors. We note that a negative slope for 'contacts' and a positive slope for 'population' support a negative coefficient for Lloyd's mean crowding.

Daily temperature (°F), relative humidity (%) and atmospheric pressure (Pa) at the centroid of each prefecture was provided by The Dark Sky Company via the Dark Sky API and aggregated across a variety of data sources. Specific humidity (kg/kg) was then calculated using the R package `humidity`[16]. Meteorological variables for each prefecture were then averaged across the entirety of the study period.

Statistical analysis. We normalized the values of epidemic peakedness between 0 and 1 and, for all non-zero values, fit a generalized linear model of the form

$$\log(Y_j) \sim \beta_0 + \beta_1 \log(C_j) + \beta_2 \log(h_j) + \beta_3 \log(P_j) + \beta_4 \log(f_j) + \beta_5 \log(t_j)$$

where, for each prefecture j, Y is the scaled inverse Shannon entropy measure of epidemic peakedness derived from the COVID-19 time series; C is the mean number of contacts[26,46]; h is the mean specific humidity over the reporting period in kg/kg; P is the estimated population density; f is the relative change in population flows within each prefecture; and t is daily mean temperature.

Projecting epidemic peakedness in cities around the world. We selected 310 urban centers from the European Commission Global Human Settlement Urban Centre Database and their included cartographic boundaries[47]. To ensure global coverage, up to the five most populous cities in each country were selected from the 1,000 most populous urban centers recorded in the database. Population count, crowding and meteorological variables were then estimated following identical procedures used to calculate these variables in the Chinese prefectures. Weather measurements were averaged over the 2-month period starting on February 1, 2020.

The parameters from the model of epidemic peakedness predicted by humidity, crowding and population size (Supplementary Table 1, model 6) were used to estimate relative peakedness in the 310 urban centers. A full list of predicted epidemic peakedness values can be found in Supplementary Table 3.

Global human mobility data. We used the Google COVID-19 Aggregated Mobility Research Dataset, which contains anonymized relative mobility flows aggregated over users who have turned on the Location History setting, which is off by default. This is similar to the data used to show how busy certain types of places are in Google Maps, helping identify when a local business tends to be the most crowded. The mobility flux is aggregated per week, between pairs of approximately 5-km^2 cells worldwide, and for the purpose of this study aggregated for 310 cities worldwide. We calculated both mobility within each city's shapefile and mobility coming into each city. For each city, change in relative mobility was defined by $m_i = m_{il}(\text{April})/m_{ib}(\text{December})$, where m_i is defined as mobility in city i.

To produce this data set, machine learning was applied to log data to automatically segment it into semantic trips[48]. To provide strong privacy guarantees, all trips were anonymized and aggregated using a differentially private mechanism[49] to aggregate flows over time (https://policies.google.com/technologies/anonymization). This research is done on the resulting heavily aggregated and differentially private data. No individual user data were ever manually inspected; only heavily aggregated flows of large populations were handled.

All anonymized trips were processed in aggregate to extract their origin and destination location and time. For example, if users traveled from location a to location b within time interval t, the corresponding cell (a, b, t) in the tensor would be $n \pm \text{err}$, where err is Laplacian noise. The automated Laplace mechanism adds random noise drawn from a zero-mean Laplace distribution and yields (ϵ, δ)-differential privacy guarantee of $\epsilon = 0.66$ and $\delta = 2.1 \times 10^{-29}$. The parameter ϵ controls the noise intensity in terms of its variance, whereas δ represents the deviation from pure ϵ-privacy. The closer they are to zero, the stronger the privacy guarantees. Each user contributes, at most, one increment to each partition. If they go from a region a to another region b multiple times in the same week, they contribute only once to the aggregation count.

These results should be interpreted in light of several important limitations. First, the Google mobility data are limited to smartphone users who have opted in to Google's Location History feature, which is off by default. These data might not be representative of the population as whole, and, furthermore, their representativeness might vary by location. Importantly, these limited data are viewed only through the lens of differential privacy algorithms, specifically designed to protect user anonymity and obscure fine detail. Moreover, comparisons across, rather than within, locations are descriptive only because these regions can differ in substantial ways.

Simulating epidemic dynamics. We simulated a simple stochastic SIR model of infection spread on weighted networks created to represent hierarchically structured populations. Individuals were first assigned to households using the distribution of household sizes in China (data from the United Nations Population Division; mean, 3.4 individuals). Households were then assigned to 'neighborhoods' of ~100 individuals, and all neighborhood members were connected with a lower weight. A randomly chosen 10% of individuals were given 'external' connections to individuals outside the neighborhood. The total population size was $n = 1{,}000$. Simulations were run for 300 d, and averages were taken over 20 iterations. The SIR model used a per-contact transmission rate of $\beta = 0.15$ per day and recovery rate $\gamma = 0.1$ per day. For the simulations without interventions, the weights were $w_{HH} = 1$, $w_{NH} = 0.01$ and $w_{EX} = 0.001$ for the crowded prefecture and $w_{EX} = 0.0001$ for the less crowded prefecture. For the simulations with interventions, the household and neighborhood weights were the same, but we used $w_{EX} = 0.01$ for the crowded prefecture and $w_{EX} = 0.001$ for the 'sparse' prefecture. The intervention reduced the weight of all connections outside the household by 75%.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

We collated epidemiological data from publicly available data sources (news articles, press releases and published reports from public health agencies) that are described in full in ref. [22]. Epidemiological and spatial data used in this study are available via Github (https://github.com/Emergent-Epidemics/COVID_crowding). The Google COVID-19 Aggregated Mobility Research Dataset used for this study is available with permission from Google. Code and data are also available at https://zenodo.org/record/4056578#.X3IFF5NKiek.

Code availability

The code associated with the data analysis and statistics is available from https://github.com/Emergent-Epidemics/COVID_crowding. The simulation code is available from https://github.com/alsnhll/SIRNestedNetwork. Code and data are also available at https://zenodo.org/record/4056578#.X3IFF5NKiek.

References
43. Ramshaw, R. E. et al. A database of geopositioned Middle East respiratory syndrome coronavirus occurrences. *Sci. Data* **6**, 318 (2019).
44. Doxsey-Whitfield, E. et al. Taking advantage of the improved availability of census data: a first look at the gridded population of the world, version 4. *Pap. Appl. Geogr.* **1**, 226–234 (2015).
45. Reiczigel, J., Lang, Z., Rózsa, L. & Tóthmérész, B. Properties of crowding indices and statistical tools to analyse parasite crowding data. *J. Parasitol.* **91**, 245–252 (2005).
46. Wade, M. J., Fitzpatrick, C. L. & Lively, C. M. 50-year anniversary of Lloyd's 'mean crowding': ideas on patchy distributions. *J. Anim. Ecol.* **87**, 1221–1226 (2018).
47. Florczyk, A. et al. GHS-UCDB R2019A - GHS Urban Centre Database 2015, multitemporal and multidimensional attributes https://data.jrc.ec.europa.eu/dataset/53473144-b88c-44bc-b4a3-4583ed1f547e (2019).
48. Bassolas, A. et al. Hierarchical organization of urban mobility and its connection with city livability. *Nat. Commun.* **10**, 4817 (2019).
49. Wilson, R. J. et al. Differentially private SQL with bounded user contribution. Preprint at https://arxiv.org/abs/1909.01917 (2019).

Acknowledgements
The authors thank K. Cordiano for statistical assistance. We thank the members of the Open COVID-19 Data Working Group. The members of the group are listed in the Supplementary Note. B.R. acknowledges funding from Google.org. M.U.G.K. acknowledges funding from the European Commission H2020 program (MOOD project) and a Branco Weiss Fellowship. O.G.P., M.U.G.K., A.E.Z. and H.T. acknowledge funding from the Oxford Martin School. H.T. acknowledges funding from the Beijing Science and Technology Planning Project (Z201100005420010). A.L.H. and A.N. acknowledge funding from the National Institutes of Health (DP5OD019851). The funding bodies had no role in study design, data collection and analysis, preparation of the manuscript or the decision to publish. All authors saw and approved the manuscript.

Author contributions
M.U.G.K., O.G.P. and S.V.S. conceived the research. B.R., A.L.H., A.N., B.A., S.V.S. and M.U.G.K. analyzed the data. B.R. and S.V.S. analyzed the human mobility data. C.D., O.G.P., M.U.G.K. and S.V.S. interpreted the data. M.U.G.K. wrote the first draft of the manuscript. All authors contributed to the interpretation of results and manuscript writing.

Competing interests
SVS is a paid consultant with Pandefense Advisory and Booze Allen Hamilton; is on the advisory board for BioFire Diagnostics Trend Surveillance, which includes paid consulting; and holds unexercised options in Iliad Biotechnologies. These entities provided no financial support associated with this research, did not have a role in the design of this study, and did not have any role during its execution, analyses, interpretation of the data and/or decision to submit.

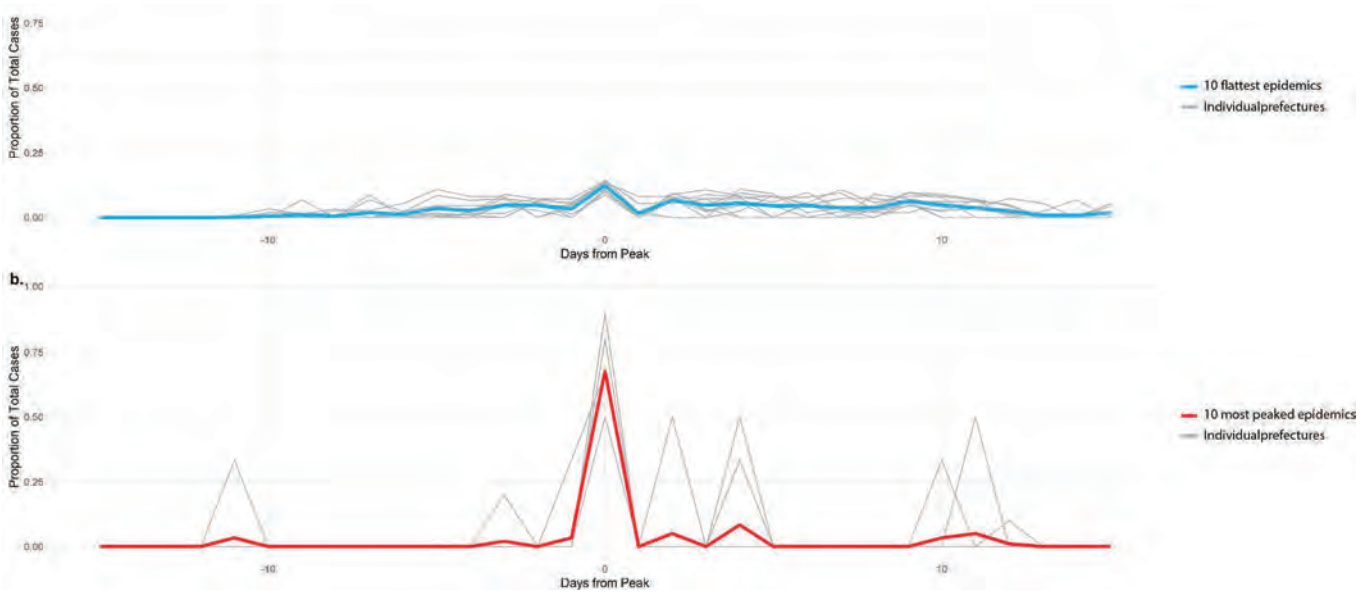
Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41591-020-1104-0.

Supplementary information is available for this paper at https://doi.org/10.1038/s41591-020-1104-0.

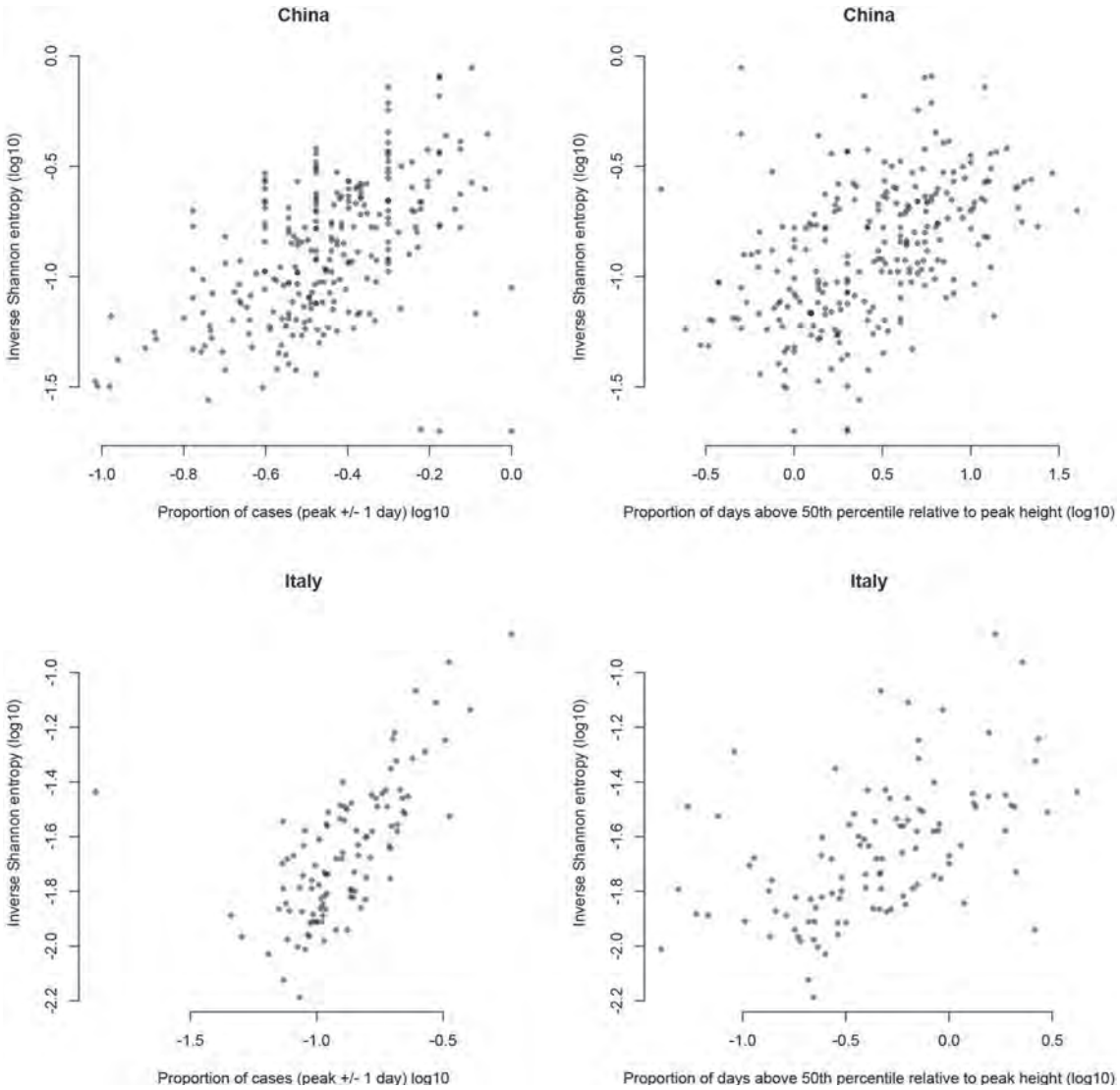
Correspondence and requests for materials should be addressed to S.V.S., O.G.P. or M.U.G.K.

Peer review information Jennifer Sargent was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

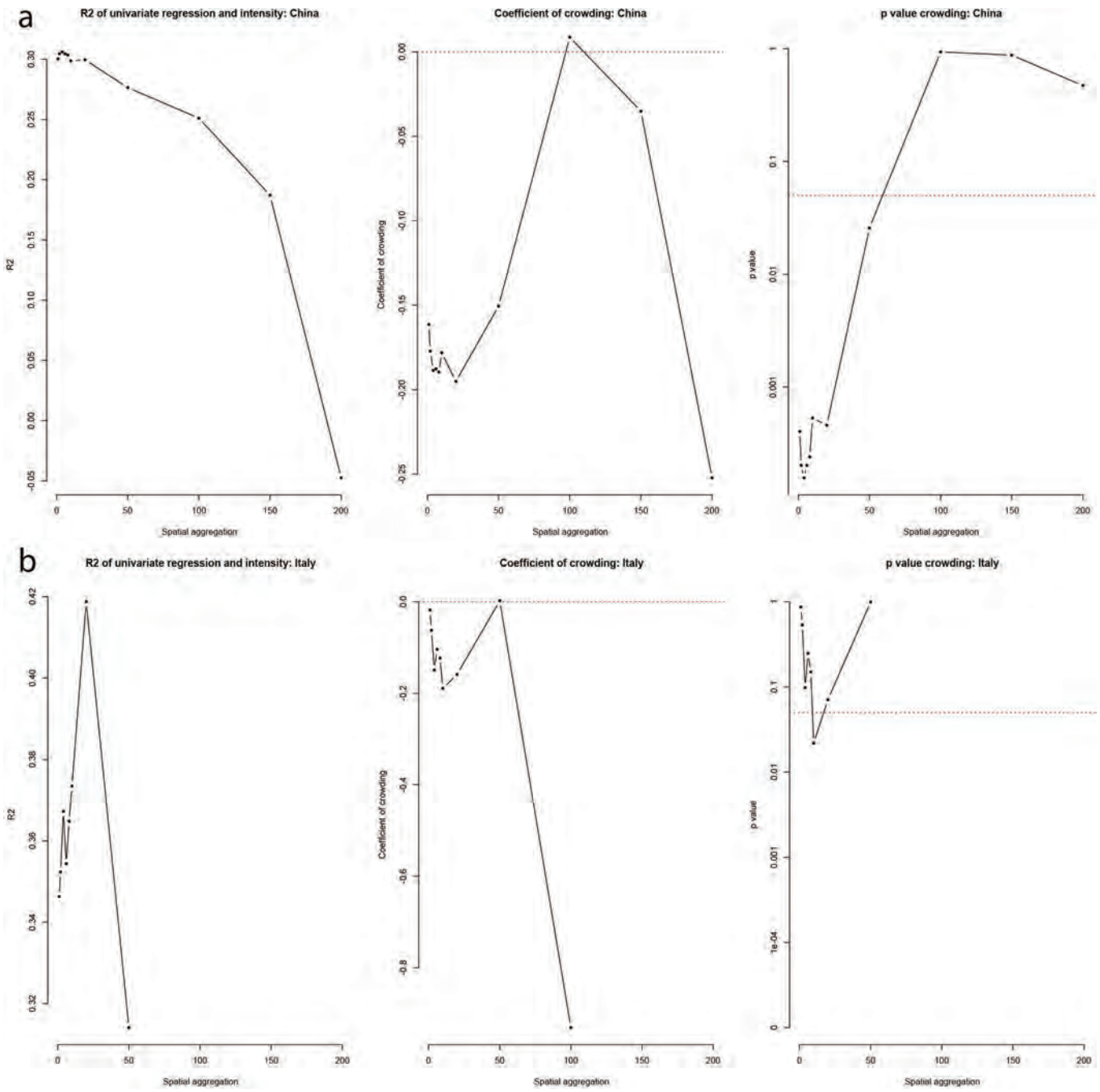
Reprints and permissions information is available at www.nature.com/reprints.



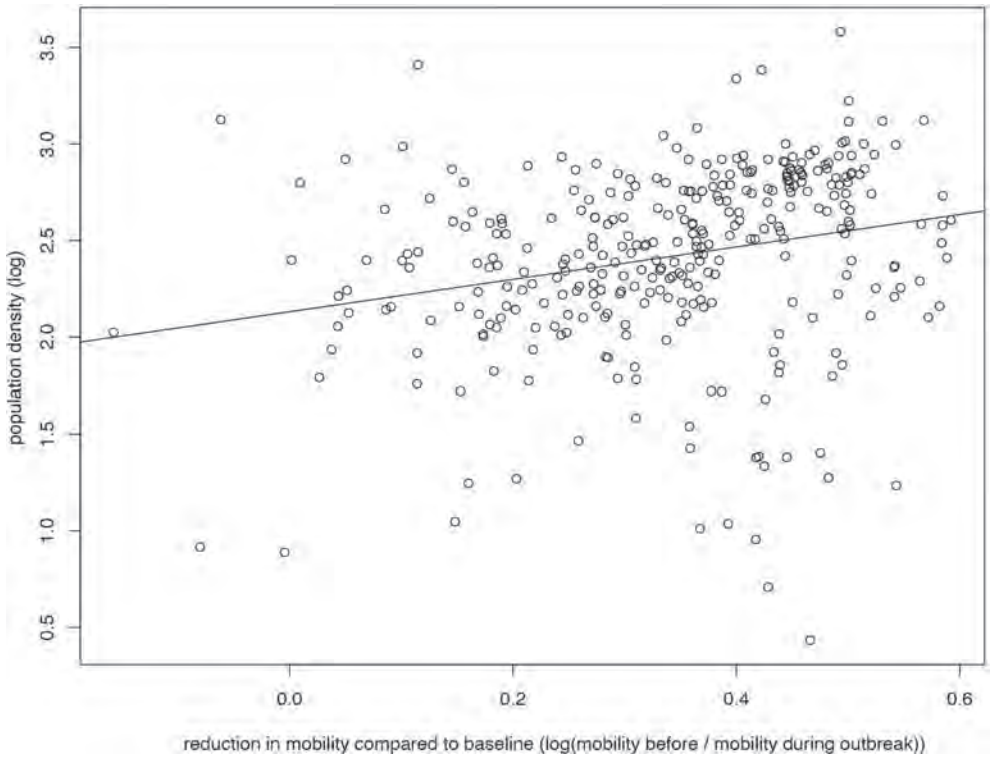
Extended Data Fig. 1 | Proportion of daily cases in prefectures in China. a, shows the ten flattest epidemics and **b**, shows the most peaked epidemics. Red and blue curves indicate the average across these prefectures.



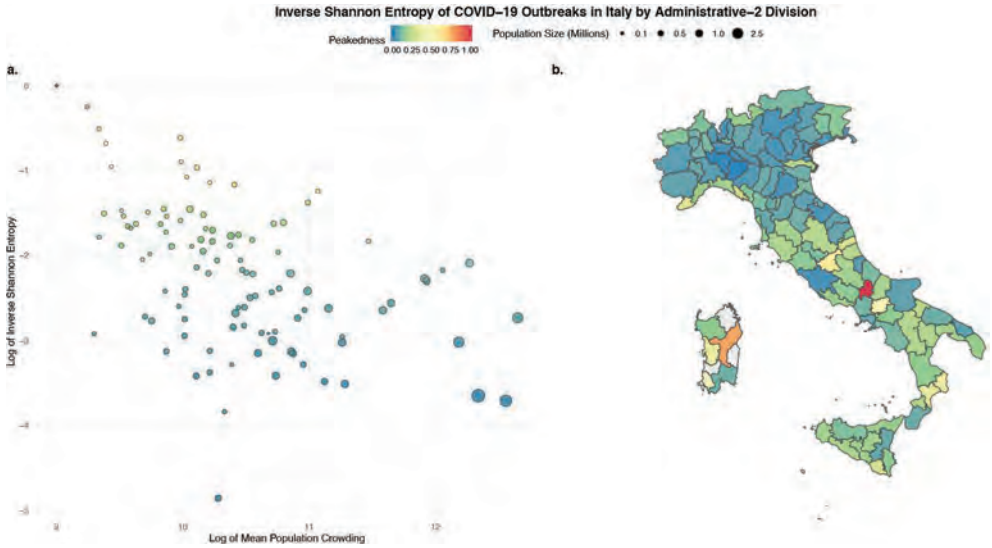
Extended Data Fig. 2 | Left panels: Proportion of cases at the peak (± 1 day) vs. inverse Shannon entropy for prefectures in China (n = 262) and regions in Italy (n = 107). Right panels: Proportion of days in the epidemic curve that had cases above the 50th percentile, normalized by the largest reported number of cases versus inverse Shannon entropy for prefectures in China (n = 262) and regions in Italy (n = 107).



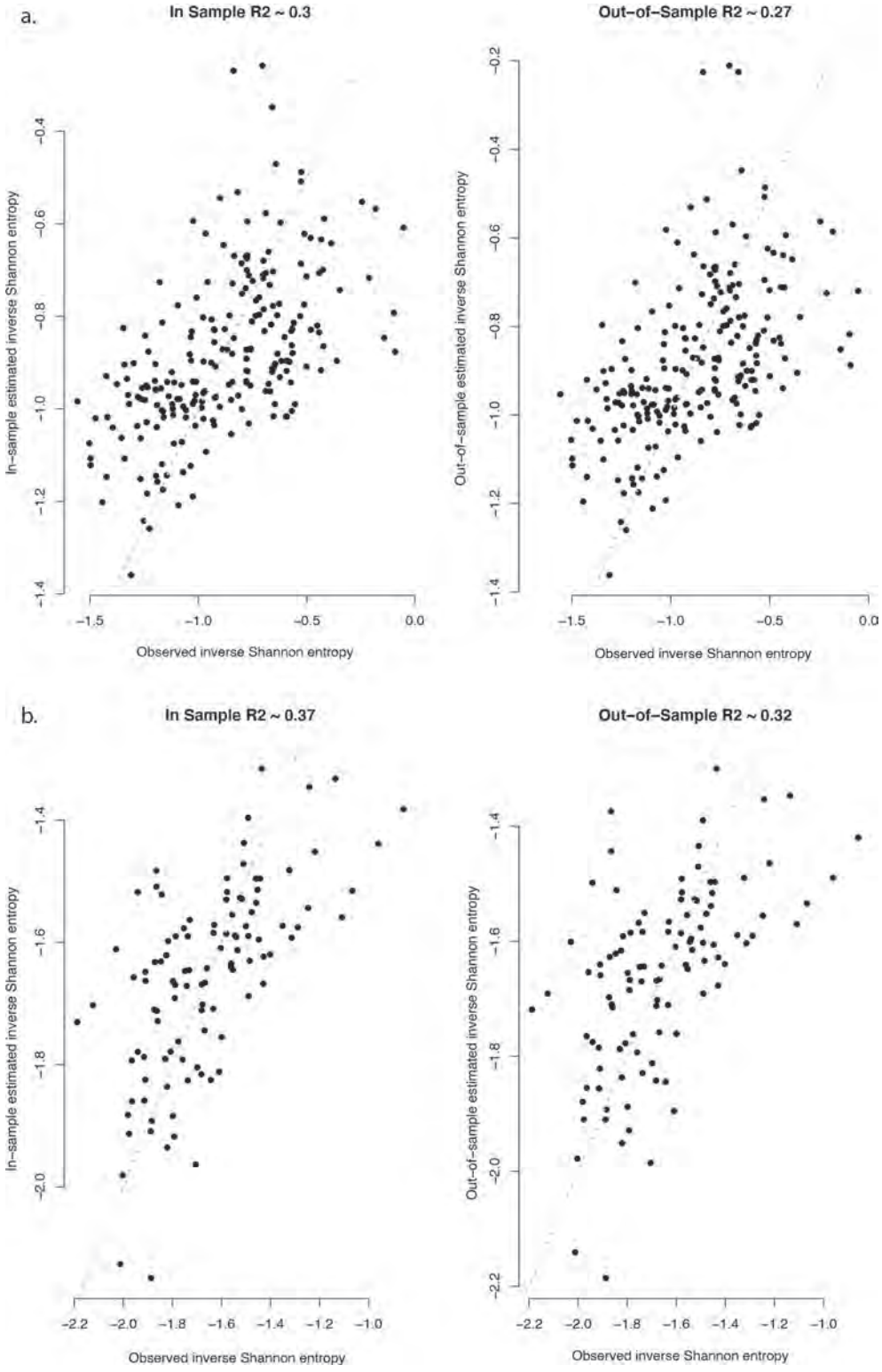
Extended Data Fig. 3 | Epidemic peakedness is well explained by covariates at spatial scales from 1 – 50 km in China **a**, and Italy **b**.



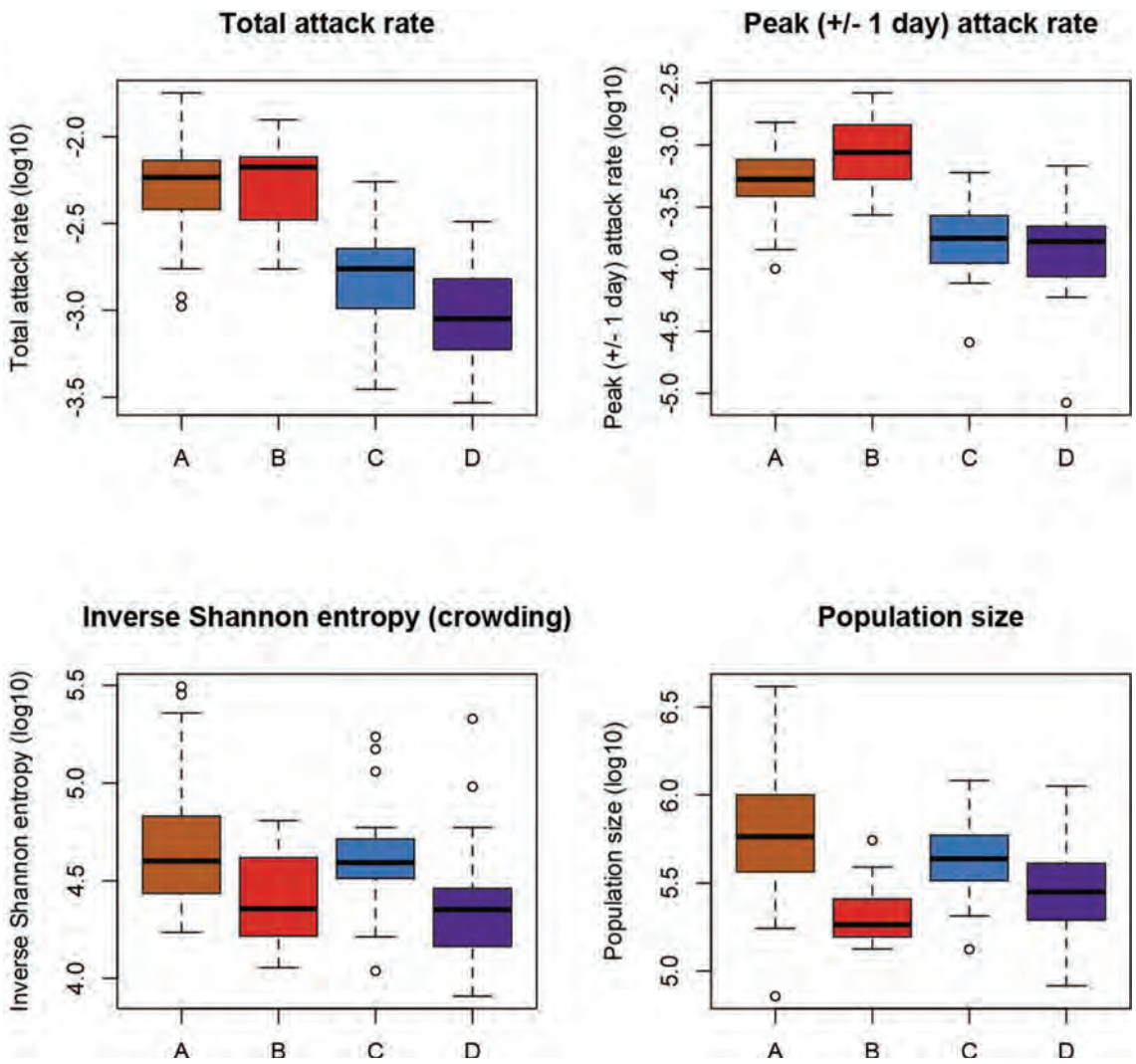
Extended Data Fig. 4 | Relationship between population density and reduction in mobility in cities in China. Each dot represents one city in China. Data on human mobility are extracted from Baidu Inc. and are available from Lai et al. 2020[36].



Extended Data Fig. 5 | Crowding and the temporal clustering of transmission of COVID-19 in Italy. a, negative association between log10 of epidemic peakedness, as measured by Shannon's diversity index (Methods), and log population crowding, as measure by Lloyd's mean crowding (Methods). The point sizes indicate the size of the population in each city, **b**, Map of epidemic peakedness in Italy at the provincal level. Blue and green colours indicate lower peakedness and red and yellow colours higher peakedness. Grey prefectures had either no or very limited amount of reported cases. Values were rescaled so that Shannon index in each province = (Shannon index—min(Shannon index)/(max(Shannon index)—min(Shannon index)).



Extended Data Fig. 6 | Out-of-sample prediction (n-fold cross validation) over all prefectures in China **a**, **b**, and all provinces in Italy **c**, **d**. In-sample prediction in China ($R^2 = 0.32$) compares well to out-of-sample predictions ($R^2 = 0.28$). In-sample prediction in Italy ($R^2 = 0.38$) compares well to out of sample predictions ($R^2 = 0.32$).



Extended Data Fig. 7 | Total attack rate, peak attack rate, inverse Shannon entropy and population sizes for: **a**, shows prefectures in China with low peakedness and high variance (as measured by the variance in the first difference of the time series of daily new cases). These prefects have high population and high crowding. **b**, shows prefectures in China with high intensity and high variance. **c**, shows prefectures in China that have low peakedness and low variance. **d**, shows prefectures in China that have high peakedness and low variance.

nature research

Corresponding author(s): Dr. Moritz Kraemer

Last updated by author(s): Jul 3, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Daily temperature (°F), relative humidity (%) and atmospheric pressure (Pa) at the centroid of each prefecture was provided by The Dark Sky Company via the Dark Sky API and aggregated across a variety of data sources (https://darksky.net/). For all other data sources no code was used.
Data analysis	R statistical software was used for the analysis: R version 3.5.1

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

We collated epidemiological data from publicly available data sources (news articles, press releases and published reports from public health agencies) which are described in full here22. Epidemiological and spatial data used in this study is available via Github (https://github.com/Emergent-Epidemics/covid_hierarchy). The Google COVID-19 Aggregated Mobility Research Dataset used for this study is available with permission from Google, LLC.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We use empirical epidemiological and spatial data to understand the intensity of transmission of COVID-19 across cities in China. Models are then used to predict epidemic intensity across a set of global cities.
Research sample	293 cities in China were used that had reported transmission of COVID-19. Prefectures in Hubei were excluded because they are distinct in their epidemiology due to the early spread of the virus there. Time series epidemiological data for Italy were available at the province level. The complete sample for data from Italy was included in this analysis.
Sampling strategy	No sampling strategy was performed. However, prefectures in Hubei were excluded. Data were sufficient in their spatial and temporal distribution for the analysis performed in this manuscript.
Data collection	Data collection is described in detail in an associated publication: https://www.nature.com/articles/s41597-020-0448-0
Timing and spatial scale	Epidemiological data for China was used from January 1 - March 30, 2020. Spatial scale was prefecture level in China (n = 293). Data from Italy was available at the prefecture level. Both datasets covered the first initial epidemic from arrival to establishment and subsequent decline in cases.
Data exclusions	Prefectures in Hubei were excluded. The epidemiological situation in Hubei is markedly different due to the early spread of the virus out of Wuhan (Hubei province).
Reproducibility	All code and data is available so that results can be reproduced.
Randomization	No randomization was performed and was not necessary for the correlative analysis performed in this work.
Blinding	No blinding was performed and necessary. Data was aggregated to daily case counts and did not include any individual level data.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Exhibit 26

Issue Brief

Fueling China's Innovation

The Chinese Academy of Sciences and Its Role in the PRC's S&T Ecosystem

Authors
Cole McFaul
Hanna Dohmen
Sam Bresnick
Emily S. Weinstein

CSET CENTER for SECURITY and EMERGING TECHNOLOGY

October 2024

Executive Summary

The Chinese Academy of Sciences is one of the most important scientific research organizations not only in China but also globally. Through its network of research institutes, universities, companies, and think tanks, CAS is a core component of China's science and technology innovation ecosystem. This brief first traces the organization's historical significance in China's S&T development, outlining key reforms that continue to shape the institution today. It then details CAS's core functions in advancing S&T research, fostering commercialization of critical and emerging technologies, and contributing to S&T policymaking. Using scholarly literature, we provide insights into CAS's research output in the science, technology, engineering, and mathematics (STEM) fields as well as in certain critical and emerging technologies, including artificial intelligence (AI).

Our key takeaways are as follows:

Research

- CAS is the top producer of STEM research globally, both in terms of total number of papers and number of highly cited papers.
- CAS is a hub of top-tier S&T researchers and plays an important role in training the next generation of experts. In 2022, CAS ranked second globally among institutions by number of top-cited researchers, trailing only Harvard University.
- CAS's 115 research institutes work on a diverse range of S&T subjects. CAS institutes publish most frequently in the field of industrial technology, which accounted for 35 percent of their published papers in 2021.
- CAS institutes also advance research in critical and emerging technologies, such as AI. The majority of AI-related research published by CAS is conducted by a narrow subset of institutes. The Institute of Automation, the Shenyang Institute of Automation, and the Institute of Computing Technology are top producers of AI-related research within CAS.

Commercialization

- CAS fosters technology transfer from research organizations to industry through various commercialization mechanisms, including making investments via asset management companies and venture capital firms, licensing proprietary research, and offering contract research services.
- CAS provides research, financing, and personnel to support the founding and development of technology companies.
- Notable companies founded with CAS support include AI company iFLYTEK, PC manufacturer Lenovo, supercomputer company Sugon, AI chip developer Cambricon, and CPU designer Loongson.

Policymaking

- CAS plays a key role in the development and implementation of China's S&T policies. CAS has contributed to major S&T policy initiatives such as the founding of the National Natural Science Foundation of China and to S&T development projects such as the 863 Program and the 973 Program.
- CAS academicians are among the most important individuals in China's S&T ecosystem. They influence resource allocation to S&T projects and often serve in important policymaking roles.

In sum, CAS is a global leader in STEM fields and a pillar of China's S&T development ecosystem. It has facilitated the rise of globally competitive technology companies and continues to influence China's S&T policy. At the same time, CAS's expansive set of responsibilities as a research organization, a commercial entity, and a bureaucratic actor complicate its mission. Understanding the tensions that may exist among these functions is important, as Beijing continues to commit extensive resources to CAS. The organization's success—or lack thereof—will affect Beijing's own ability to effectively achieve its S&T development ambitions.

Table of Contents

Introduction

The Chinese Academy of Sciences (中国科学院) is a key actor in China's science and technology ecosystem. Directly managed by the State Council, CAS plays a critical role in advancing Beijing's S&T ambitions and is deeply involved in shaping and implementing national policies.[1] As technological innovation is a core aspect of the intensifying competition between the United States and China, it is imperative that U.S. policymakers understand CAS's contributions to China's S&T development.

In this brief, we seek to advance a better understanding of CAS, which is one of the largest and most prolific research bodies in the world and home to many top science, technology, engineering, and mathematics (STEM) researchers.[2] CAS oversees hundreds of subsidiary organizations, including 115 research units,* three universities, numerous companies, and several think tanks. Comprising an expansive network of researchers, students, and policymakers, CAS directly contributes to both the design and implementation of Chinese S&T policies.

This brief proceeds as follows. First, we describe CAS's history and trace successive reforms to the institution, highlighting its role in China's rise as a major S&T power. We then detail CAS's core functions in China's S&T ecosystem: advancing research, promoting the commercialization of key technologies, and contributing to policymaking. We conclude with an examination of CAS's central position in China's S&T ecosystem and its implications for U.S. policymakers.

* Throughout this piece, we refer to CAS's research units (研究单位) as CAS institutes.

Methodology

Our analysis draws on a variety of publicly available sources, including Chinese government policy documents, annual reports of publicly listed companies, and an array of other primary and secondary sources. Our analysis of CAS's research output relies on two distinct data sources to present trends in CAS's English- and Chinese-language STEM research.

First, we draw from CSET's merged corpus of scholarly literature published between 2010 and 2023, which includes journal articles, also referred to in this brief as papers, from Clarivate's Web of Science, Semantic Scholar, the Lens, OpenAlex, arXiv, and Papers With Code. CSET's merged corpus primarily consists of English-language journal articles. We use CSET's fields of study methodology to classify papers by field.[3] We focus on CAS's STEM research because one of the organization's core responsibilities is advancing China's S&T development and self-reliance.[*] We define STEM publications as papers in the fields of biology, physics, geology, mathematics, chemistry, computer science, engineering, environmental science, and materials science.

Second, we leverage data from the China National Knowledge Infrastructure, which is largely composed of Chinese-language papers, to analyze CAS's STEM publications within China's domestic research ecosystem. For this report, we draw on CNKI papers published between 2010 and 2021.[†] For CNKI papers, we define STEM publications as those in the fields of industrial technology, astronomy and geoscience, agricultural sciences, mathematical sciences and chemistry, environmental science, life sciences, aviation and aerospace, and transportation.[‡] Finally, we use the CNKI dataset to classify AI-related publications in order to better understand CAS's contributions to China's AI development ecosystem.[4]

By analyzing both CSET's merged corpus and CNKI, we are able to conduct a more comprehensive analysis of CAS's research output. Using the merged corpus, we assess CAS's research output alongside that of other global research institutions. CNKI, on the other hand, includes more detailed paper publication information for each subsidiary

[*] We are also particularly interested in CAS's contributions to AI-related research and development, given its strategic importance to Beijing and relevance in U.S.-China tech competition.

[†] We are unable to analyze CNKI publication data in 2022 and 2023 due to incomplete data.

[‡] CNKI employs the Chinese Library Classification system, which categorizes publications by subject field. Every paper in CNKI receives a CLC code that corresponds to a specific field of study. The CLC fields of study are not the same as those used in CSET's merged corpus.

organization within CAS, which allows us to analyze the publications of specific entities, such as its 115 research institutes.

That said, our methodological approach has some limitations. First, our dataset of CAS research is not exhaustive. Within both CSET's merged corpus and CNKI, CAS research is often attributed to the organization's subsidiaries, rather than to CAS itself. As a result, querying the datasets for publications strictly attributed to CAS would miss a significant portion of CAS-affiliated research. To address this issue, we use two approaches. First, CSET's entity resolution methodology helps to match subsidiaries to their parent organizations in the merged corpus dataset. Second, we use CNKI-provided organizational information to carefully identify all the publications of each CAS institute, which is not possible using the merged corpus data. This methodology enables a more comprehensive assessment of CAS's research. At the same time, our methodology does not capture all publications of every CAS subsidiary organization. For example, our analysis of the CAS institutes' publications does not account for some CAS-authored research papers published by other CAS subsidiary organizations, such as universities and laboratories.

Second, we use several proxies for evaluating research quality, such as highly cited research publications, top-cited research publications, and number of top-cited researchers.* Although imperfect, these metrics allow us to compare measures of high-impact research and researchers for CAS with those of other international research institutions. Importantly, these metrics do not provide insight into the production of highly cited research relative to size, research expenditures, or other relevant factors that may contribute to the output of research institutions. While important to consider, these cross-institutional analyses are largely outside the scope of this report.

* We define highly cited papers as papers in at least the 90th percentile of citations in their field in a given year, top-cited papers as papers in the 99th percentile of citations in their field in a given year, and top-cited researchers as in the 99th percentile of individual researchers most frequently cited in their field and year.

History, Development, and Reform

Established in November 1949, CAS has played an instrumental role in laying the foundation for China's modern S&T ecosystem. Tracing CAS's history of development and reform is essential to understanding its position in Chinese S&T development today.

The Early Days: Defense S&T and Building on the Soviet Model

In the 1950s, the Academy of Sciences of the Soviet Union strongly influenced CAS's development.[5] Collaboration with the Soviet Union helped propel China's strategic weapons development programs. In February 1953, for example, a delegation of Chinese scientists led by the director of the Institute of Physics of CAS traveled to the Soviet Union to discuss atomic research collaboration.[6] In 1957, CAS and the Soviet Academy of Sciences signed joint agreements to construct nuclear reactors in Chongqing, Xi'an, Beijing, and Shenyang.[7]

CAS played a critical role in China's strategic weapons and space technologies development throughout the 1950s and 1960s. At the instruction of Premier Zhou Enlai (周恩来), CAS made significant contributions to China's "Nuclear Bombs, Ballistic Missiles, and Earth Satellites" (两弹一星) program, which led to the development of the country's nuclear and space capabilities.[8]

Another example of CAS's success in supporting Chinese military modernization occurred in March 1965, when the Chinese Communist Party tasked CAS with developing a guidance computer for long-range missiles, called Project 156.[9] The consortium of six CAS institutes working on the project formed the Xi'an Microelectronics Technology Institute (also known as the 771 Institute; 西安微电子技术研究所) in 1975, which developed into an important supplier of computing and microelectronics products to the People's Liberation Army.[10] Notably, ZTE (中兴通讯股份有限公司), one of the largest telecommunications companies in the world, was spun out of the Xi'an Microelectronics Technology Institute in the early 1980s.[11]

Reform and Revitalization

In the 1970s and 1980s, CAS suffered from an aging workforce, insufficient research facilities, and outdated research agendas, all of which hampered its ability to advance indigenous innovation and other S&T development priorities.[12] During the 1980s, national reforms decreased government funding for research institutions and incentivized the organization to become more responsive to market competition. These reforms discouraged CAS entities from undertaking large-scale research projects with uncertain commercial outcomes and ultimately led to the underfunding of CAS's basic research activities.[13]

In the late 1990s, responding to Beijing's calls to strengthen China's national research capabilities, CAS initiated a series of efforts to expand its role in China's innovation ecosystem.[14] The most significant of these efforts was the Knowledge Innovation Program (中国科学院知识创新工程), a set of initiatives and reforms intended to address fundamental issues within CAS and develop it into one of the world's top scientific research institutions.[15]

The KIP reduced the number of CAS institutes by converting some into commercial entities and reorganized others to address overlapping research missions.[16] The KIP included several initiatives to bolster CAS's research workforce, such as expanding training for graduate students, introducing stringent evaluation requirements for researchers, and leveraging national-level initiatives like the Hundred Talents Program (百人计划) to recruit top scientists from abroad.[17] Later, the KIP encouraged collaborations between CAS and provincial and city governments to support local innovation, leading to the creation of seven new institutes.[18] By the end of the 12-year program, CAS reemerged as a hub of basic research and S&T development in China.[19]

In 2011, CAS unveiled a new initiative called Innovation 2020 (创新 2020), which aimed to continue the work of the KIP and further strengthen China's S&T ecosystem. Innovation 2020 focused on improving international collaborations, including through deepening scientific partnerships with developed nations and promoting cooperation with developing countries.[20] As part of the KIP, CAS played an active role in China's Belt and Road Initiative. For example, between 2013 and 2019, it provided around 268 million USD in funding for BRI S&T projects and initiated collaborative research with over 40 institutions from 14 BRI partner countries.[21]

CAS's 13th Five-Year Plan (2016–2020) placed greater emphasis on converting basic research into commercial technologies that can help drive economic growth.[22] During this period, CAS established joint research and development (R&D) centers with major

industry players, set up business incubators for startups, improved patenting processes and standards, and further promoted collaboration with local governments. It also reformed its performance assessment system—which directly affects resource allocations—to place a greater emphasis on conducting impactful, high-quality research, replacing its previous focus on using quantitative metrics to evaluate impact.[23]

Since the 1950s, CAS has been a key driver of S&T development in China. Successive reforms over the past four decades have led to CAS's modernization and have better aligned incentive structures to advance domestic S&T innovation. These reforms have also promoted CAS's integration into the international scientific landscape. In the following section, we build on this context to outline CAS's contemporary contributions to China's S&T ecosystem.

Examining CAS's Core Functions: Research, Commercialization, and Policymaking

In this section, we examine CAS's three main functions: advancing research in critical fields, fostering the commercialization and adoption of key technologies, and aiding in the development and implementation of China's S&T policy initiatives. In the first section on CAS's role in driving research, we detail the organization's size, funding, and research output in the context of other leading global research institutions. We also include analysis of the CAS institutes' research activities by field of study to better understand CAS's research output in the context of China's STEM ecosystem, paying special attention to the institutes' AI-related publications. The second section describes CAS's role in fostering the commercialization of key technologies and highlights several key case studies of successful CAS spin-off companies. The third section describes CAS's role in shaping Chinese S&T policy.

Advancing S&T Research

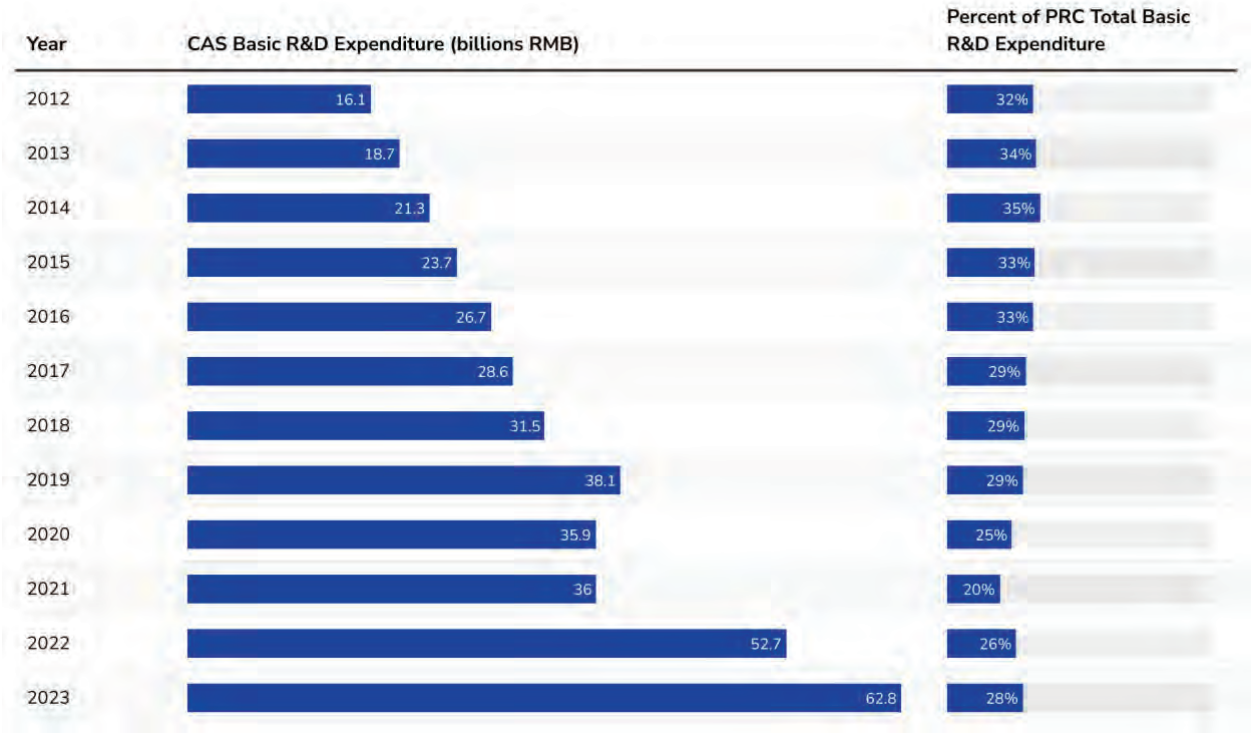
One of CAS's primary responsibilities is to advance China's S&T research capabilities in order to strengthen the national innovation ecosystem, boost technological self-reliance, and help China become a global leader in STEM research.[24] CAS oversees 115 institutes, which advance research across a wide range of critical fields. CAS also directly manages two universities, the University of Science and Technology of China (USTC) and the University of Chinese Academy of Sciences (UCAS), and co-administers ShanghaiTech University with the Shanghai Municipal People's Government.[25] Additionally, CAS is an important actor in China's State Key Laboratory system, which is composed of hundreds of research facilities tasked with conducting cutting-edge research and fostering cooperation among universities, research institutes, and companies.[26] CAS oversees 83 (roughly 30 percent) of China's government-managed SKLs, many of which are co-located with CAS institutes and universities.[27] Finally, CAS sponsors 267 academic journals, some of which are published by Springer and Oxford Academic.[28]

Taken together, these universities, research institutes, and SKLs form a network of research entities that work in concert to advance scientific progress and bolster China's STEM workforce. CAS universities, for example, work with CAS institutes and CAS SKLs to provide students with opportunities to contribute to research projects. Some scientists at CAS institutes and SKLs serve as part-time department leaders at CAS universities or otherwise contribute to CAS's talent training system.[29] According to CAS, it employs around 71,000 individuals, almost 62,000 of whom are professional researchers. Around 79,000 graduate students study at CAS and its universities.[30]

CAS is an important hub for basic research in China. CAS-affiliated researchers account for 40 percent of all principal investigators overseeing scientific research funded by the National Natural Science Foundation of China (NSFC; 国家自然科学基金委员会).[31] As of May 2016, CAS oversaw 80 percent of China's large-scale science facilities, including the China Spallation Neutron Source in Dongguan, which supports research on physics and materials science, among other fields, and the Experimental Advanced Superconducting Tokamak in Hefei, which facilitates research on nuclear fusion technology.[32]

Figure 1 shows CAS's basic R&D expenditures between 2012 and 2023, collected from CAS's annual budget reports. CAS's basic R&D expenditures have nearly quadrupled since 2012 and have grown markedly in the last few years, increasing from 36 billion RMB (5.6 billion USD) in 2021 to 62.8 billion RMB (8.9 billion USD) in 2023.[33]

Figure 1. CAS Basic R&D Expenditure and Percent of China's Total Basic R&D Expenditure, 2012-2023



Source: CAS Budget (2012–2023), National Bureau of Statistics.[34]

Figure 1 underscores CAS's consequential role in China's basic research ecosystem. As shown above, CAS's basic R&D expenditures accounted for 28 percent of all such spending in China in 2023. While this proportion has declined from its 2014 peak, CAS remains a key actor in China's basic research landscape.

At the same time, these figures should be understood within the broader context of China's R&D environment, which continues to favor expenditures on later-stage projects over basic research. Moreover, China's proportional spending on basic R&D continues to trail that of other leading S&T countries. According to a 2018 U.S. National Science Foundation study, for example, China's basic R&D expenditure accounted for just 5 percent of total R&D spending, the lowest of any surveyed country.[35] In the United States, basic R&D spending accounted for 16 percent of all R&D expenditures in 2018.[36]

In short, CAS is a major player in China's R&D efforts, driving S&T advancements through its vast network of institutes, universities, and SKLs. The organization also contributes to the training of China's next generation of scientists, oversees major basic

research projects, and manages large-scale scientific facilities. To provide further insight into CAS's research contributions, we examine its research output in STEM fields in the following section.

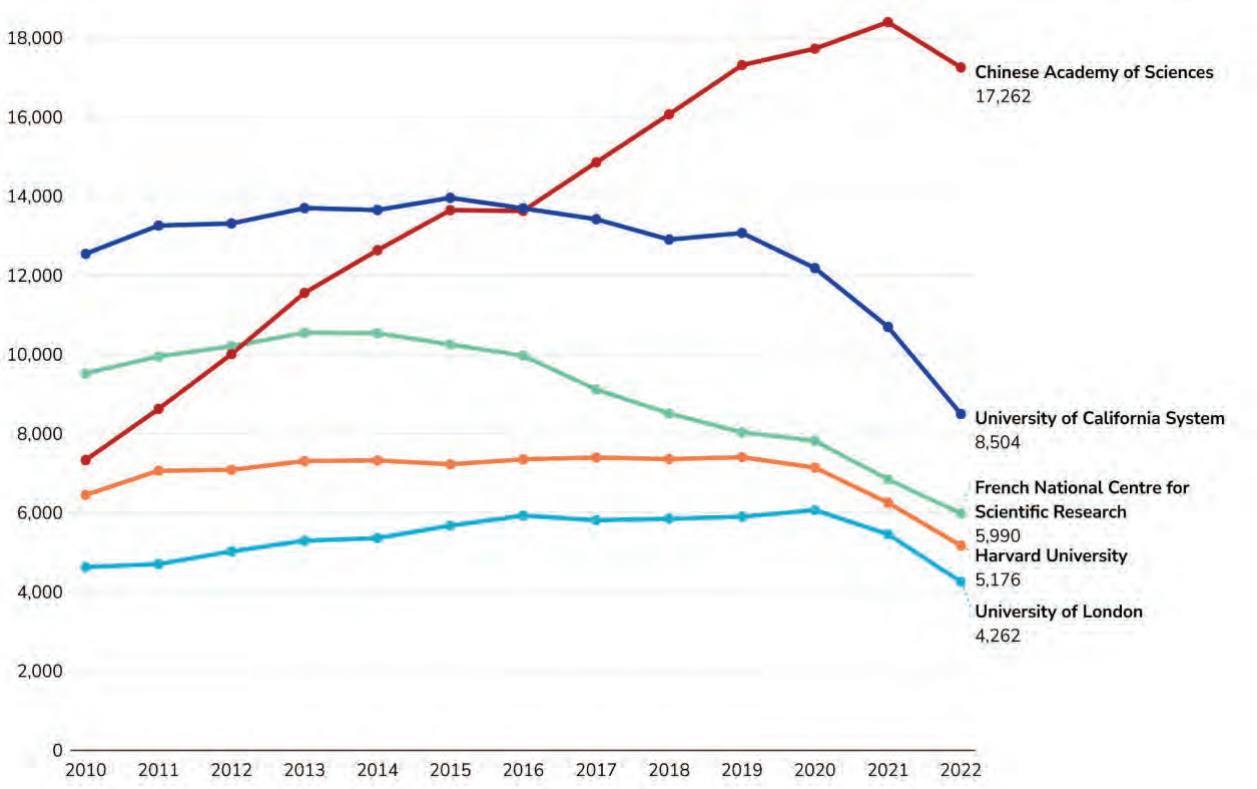
CAS Research on the Global Stage

CAS is the world's top producer of STEM research in terms of volume of papers produced, ahead of the French National Centre for Scientific Research and the University of California system (see Figure A1 in Appendix A), according to our analysis of CSET's merged corpus data. Since 2012, CAS has also topped the annual Nature Index, which ranks the top five hundred global research institutions based on counts of papers published in top international natural science and health science journals.[37] In 2023, UCAS and USTC—the two universities managed by CAS—ranked fifth and seventh on the Nature Index, respectively.[38]

Even when only accounting for highly cited STEM papers, CAS tops all other global research institutions.* We focus on highly cited papers, as previous bibliometric research has found that citations can be an indicator, albeit an imperfect one, of relatively high-impact research.[39] As previous research has shown, Chinese researchers are strongly incentivized to write and publish frequently.[40] This dynamic has led to a proliferation of low-quality papers published, sometimes by so-called paper mills that eschew rigorous reviews in favor of charging authors fees in order to publish papers.[41] Yet, as shown in Figure 2, CAS researchers are producing an increasing number of highly cited papers relative to other global research institutions.

* We define highly cited papers as papers in at least the 90th percentile of citations in their field in a given year.

Figure 2. Top Research Institutions by Highly Cited STEM Publications, 2010–2022



Source: CSET Merged Corpus.

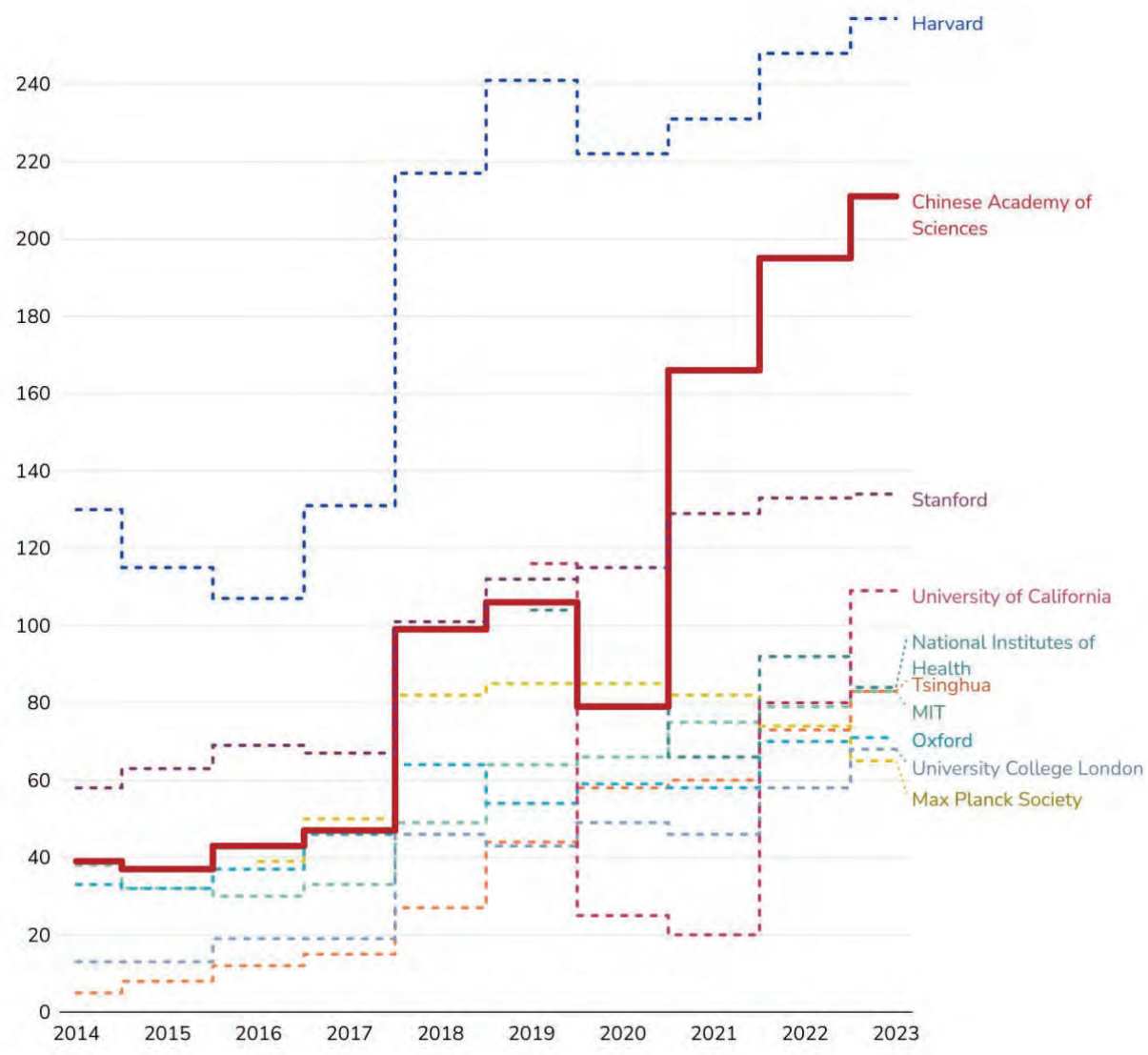
Note: See Figure A2 in Appendix A for the top 15 global research organizations for counts of both total and highly cited STEM papers.

Between 2010 and 2022, CAS's highly cited STEM research output more than doubled, its growth outpacing other leading global hubs of STEM research. Highly cited STEM papers from the University of California system and Harvard University increased only modestly over the same period, and yearly publications of highly cited papers from the other four institutions shown in Figure 2 have declined since 2020, likely due in part to the COVID-19 pandemic. In contrast, the number of CAS's highly cited STEM research publications has not declined nearly as much. Accounting only for top-cited STEM papers, our research shows that CAS passed the University of California system, the largest university system in the United States, in 2018.* Figure A3 in Appendix A shows the growth in top-cited STEM papers over time among the leading five producers of top-cited research in the world.

* We define top-cited papers as papers in the 99th percentile of citations in their field in a given year.

Not only is CAS a leading global research institution in terms of highly cited and top-cited STEM research, but individual CAS researchers are often among the leaders in their fields. Since 2014, Clarivate has published an annual list of top-cited researchers, which tracks the affiliations of individuals that rank within the top 1 percent of the most frequently cited researchers in their field and year.[42] Clarivate's list of top-cited researchers includes both STEM and non-STEM fields. As Figure 3 shows, the number of top-cited researchers affiliated with CAS has increased dramatically over the past decade, growing from 39 in 2014 to 211 in 2023, trailing only the number of those at Harvard University.

Figure 3. Highly Cited Researchers by Institution, 2014–2023



Source: Clarivate.[43]

Note: In earlier years, highly cited researchers were not aggregated for the University of California system (aggregated after 2018), the National Institutes of Health (aggregated after 2018), and the Max Planck Society (aggregated after 2015).

To be sure, readers should be judicious when drawing conclusions from the above figures, as the featured institutions are not immediately comparable without additional context. Among other differences, the size of these institutions varies, both in terms of research expenditures and number of researchers. Harvard University and Stanford University, for example, both spent around 1.3 billion USD on R&D in 2022, compared to CAS's more than 15 billion USD.[44] As of 2019, Harvard reported 7,579 R&D

personnel, much smaller than the 56,000 reported by CAS.[45] On the other hand, in both total operating budget and number of faculty and staff, CAS is smaller than the University of California system.[46] In short, while our data allows us to assess measures of STEM research output, we do not offer analysis of leading research institutions' publications relative to other important factors, including R&D expenditures and number of R&D personnel.

Still, the data presented above suggests CAS is publishing not only more STEM research papers but also a growing share of the world's most highly cited research. Over the last decade, the number of highly cited CAS publications and researchers has increased substantially, representing a rise in impactful research. CAS institutes have played an especially important role in fostering CAS's growth in high-quality research. We further detail the function and research output of the CAS institutes below.

CAS Institutes

CAS institutes are tasked with leading China's basic and applied research to solve major strategic and technological challenges, boost China's S&T self-reliance efforts, and make China's S&T ecosystem more internationally competitive.[47] The 115 CAS institutes are spread across 25 provinces and municipalities, with 35 in Beijing, 15 in Shanghai, and seven in both Jiangsu and Guangdong (see Appendix B for a complete list of CAS institutes). Below, we turn our attention to the role CAS institutes play in advancing STEM research within China's S&T ecosystem and provide some quantitative measures of their research output. We pay particular attention to the institutes' research output in critical and emerging technology fields, including in AI, to better understand CAS's efforts to advance China's capabilities in technologies deemed strategically important by Beijing.

While each institute is responsible for its own research and administration, CAS headquarters oversees the CAS institute system. CAS appoints each institute's leadership, authorizes research strategies, allocates resources, and assesses research output quality and performance. All CAS institutes are expected to adhere to the "One-Three-Five" guideline, which instructs them to have one R&D direction, to make three major breakthroughs within five to 10 years, and to establish five key line-of-research priorities every five years.[48] As a result of this organizational structure, each CAS institute specializes in specific scientific areas.

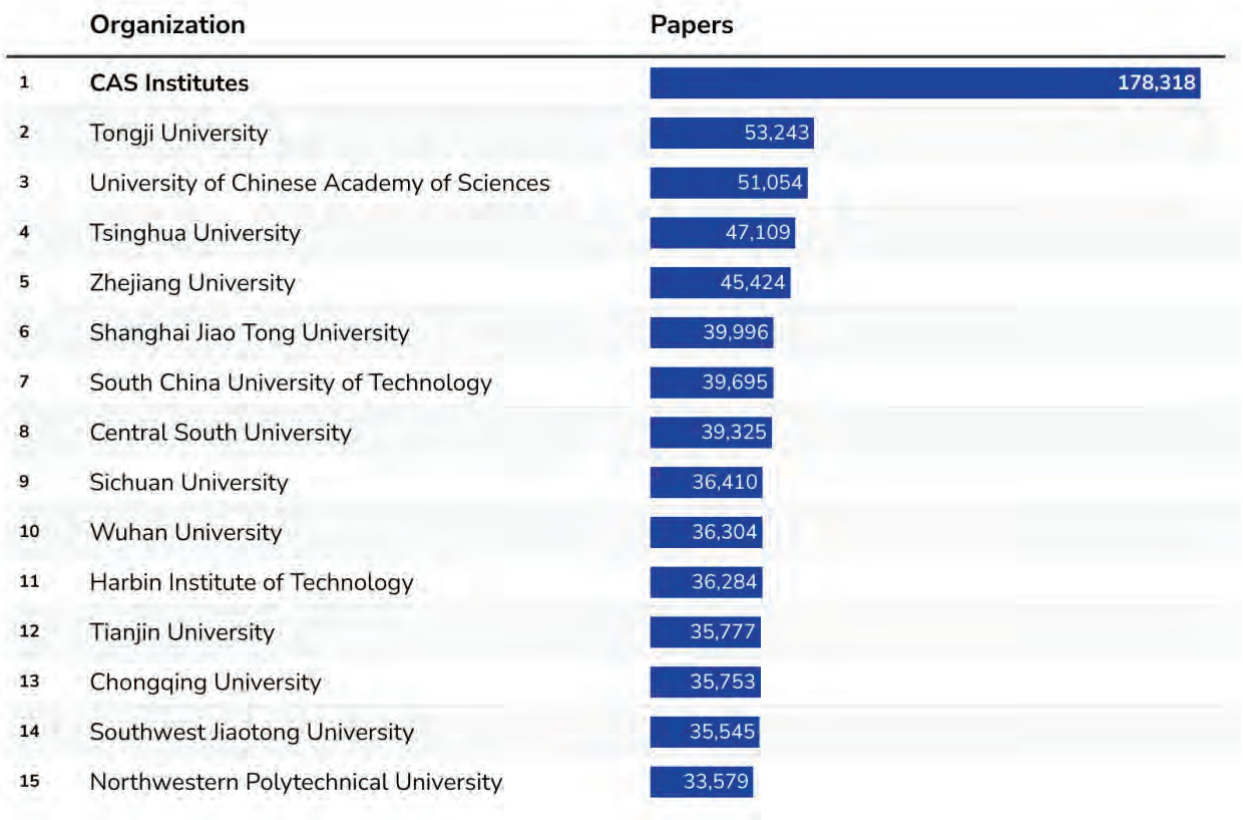
In addition to their role in advancing S&T research, CAS institutes help develop China's S&T workforce by providing graduate students with training and hands-on research opportunities. For example, UCAS offers a "two-stage" (两段式) program, which

allows graduate students to first take courses at the university and then complete degree-related scientific research at one of the CAS institutes.[49] Moreover, some faculty hold joint appointments at a CAS university and a CAS institute, helping to bring additional scientific expertise from practicing scientists directly into the classroom.[50] Several CAS institutes offer graduate degree programs, sometimes in partnership with neighboring universities.[51] These mechanisms provide graduate students enrolled at CAS universities with valuable research experience and access to scientific research facilities.

Below, we draw on CNKI data to assess the 115 CAS institutes' publications by field of study. Each scientific paper in CSET's repository of CNKI data includes information about the authors and their organizational affiliations. Some CAS-affiliated authors may list their affiliation to a certain CAS institute but not to CAS itself. By assessing the publications attributed to CAS institutes, we are able to capture a more comprehensive sample of CAS-affiliated research within China's S&T research ecosystem.

In Figure 4, we present the STEM papers of the CAS institutes alongside those of the top fifteen research organizations that appear in CSET's CNKI dataset. Taken together, CAS institutes are the most prolific producers of STEM research in China, demonstrating the importance of CAS-affiliated research organizations in China's S&T ecosystem. Although the research institutions in Figure 4 are not directly comparable, as they each vary in size and structure, it is nonetheless noteworthy that the CAS institutes produced more than three times as many STEM publications as the next leading Chinese research institution. The CAS institutes accounted for 178,318 papers published between 2010 and 2021, over 125,000 more than Tongji University, the second largest producer of STEM research in China. UCAS—the largest of the three CAS-managed universities—ranked third.

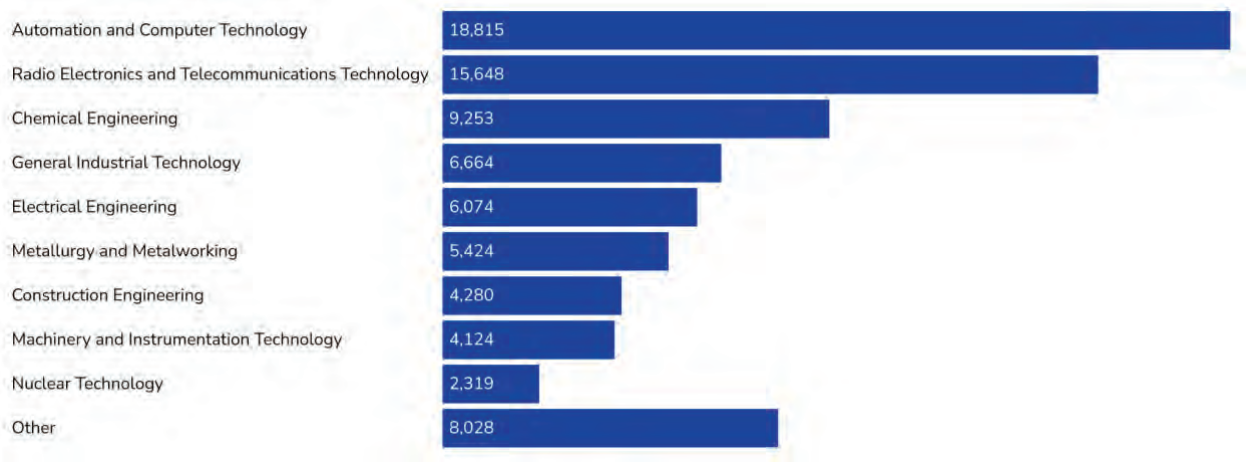
Figure 4. STEM Paper Publications by Top Publishers in China, 2010–2021



Source: CNKI.

The research output of the CAS institutes encompasses a wide range of disciplines. Reflecting CAS's focus on research in critical and emerging technology fields, CAS institutes published most often in the industrial technology field, which encompasses technologies related to autonomy, information and communication technologies, chemical engineering, and more. Industrial technology accounted for 35 percent of CAS institute papers in 2021, followed by astronomy and geoscience, environmental science, and agricultural sciences. See Figure A4 in Appendix A for CAS institute papers over time by field. Figure 5 shows CAS industrial technology papers by subfield. CAS institutes published most frequently in the subfield of automation and computer technology, which includes AI-related research, followed by radio electronics and telecommunications technology and chemical engineering.

Figure 5. CAS Institutes Industrial Technology Papers by Subfield, 2010–2021



Source: CNKI.

CAS is especially important to China's AI research. According to *China's New Generation Artificial Intelligence Technology Industry Development Report 2022*, published by the Chinese Institute of New Generation Artificial Intelligence Development Strategies, CAS is a key node within China's AI innovation network and an important source of basic research for China's top AI companies.[52] Our analysis of the research output of the CAS institutes further supports CAS's role in China's AI ecosystem. Between 2010 and 2021, the CAS institutes published 23,431 AI-related papers in STEM fields, more than twice as many as Wuhan University, the second leading Chinese institution in AI-related publications. See Figure A5 in Appendix A for the top ten publishers of AI-related publications in China.

Some CAS institutes are focused on advancing AI research. The Institute of Automation (自动化研究所), the Shenyang Institute of Automation (沈阳自动化研究所), and the Institute of Computing Technology (计算技术研究所) are among the top producers of AI-related research within CAS. Over half of all papers published by these institutes were AI-related, as shown in Figure 6.

Figure 6. CAS Institutes by Percentage of AI-Related Papers, 2010–2021

CAS Institute	AI Papers	Total Papers	AI Papers, Percent of Total
Institute of Automation	1,380	1,635	84%
Shenyang Institute of Automation	1,963	2,744	72%
Institute of Computing Technology	1,615	2,737	59%
Aerospace Information Research Institute	417	863	48%
Institute of Software	781	1,770	44%
Institute of Information Engineering	614	1,463	42%
Suzhou Institute of Biomedical Engineering and Technology	166	534	31%
Institute of Optics and Electronics	426	1,495	29%
Hefei Institutes of Physical Science	429	1,525	28%
Institute of Acoustics	759	2,925	26%
Academy of Mathematics and Systems Science	350	1,447	24%
Changchun Institute of Optics, Fine Mechanics, and Physics	1,984	8,266	24%
Xi'an Institute of Optics and Precision Mechanics	415	1,791	23%
National Space Science Center	238	1,041	23%
Shanghai Institute of Technical Physics	438	2,191	20%

Source: CNKI.

Note: This figure only includes CAS institutes that published at least one hundred AI-related papers between 2010 and 2021.

Notably, a small number of CAS institutes conducted the majority of CAS's total published AI-related research. Eight of them accounted for half of all AI-related papers published by CAS institutes between 2010 and 2021. This degree of specialization likely reflects their adherence to the "One-Three-Five" guideline, which encourages CAS institutes to deepen their expertise in more specialized research fields.

As shown above, CAS is a hub for top-tier S&T talent and a leader in STEM research, both on the global stage and within China's domestic innovation ecosystem. But CAS is not only a research institution. In the following section, we highlight CAS's efforts to commercialize research in key technology areas.

Fostering Commercialization of Critical Technologies

A top priority for Beijing, commercializing S&T research is another of CAS's core functions. The CAS 13th Five-Year Plan stipulates that the organization incubate five thousand "'popular entrepreneurship and mass innovation' enterprises, strengthen and expand a batch of globally competitive innovative enterprises and 'hidden champion' (隐形冠军) enterprises, and provide the 'four technology' services (technology development, technology ownership transfer, technology consulting, and technology services) for no less than 20,000 enterprises."[53] Moreover, in recent years CAS has implemented policies urging universities and research institutes to allow academics to work part-time at companies and has added technology transfer metrics to the research institution evaluation system.[54] These efforts suggest that CAS is redoubling its efforts to be a bridge between basic and applied research initiatives, as well as to promote technology diffusion and adoption.

CAS has various commercialization mechanisms at its disposal to encourage technology transfer from research institutions to industry. These include offering contract research services, licensing proprietary technology, launching new companies out of CAS institutes, providing industry access to CAS research facilities, and providing financing for companies through investment mechanisms.[55]

Chinese Academy of Sciences Holding Co., Ltd. (中国科学院控股有限公司), for instance, is CAS's primary investment arm. Established in 2002, CASH invests in five key sectors: new materials, energy, and environmental protection; technology services and financial technology; publishing and media; high-end equipment; and information technology.[56] CAS also operates a venture capital investment arm—CAS Investment Management Co., Ltd. (中国科技产业投资管理有限公司)—which primarily makes early-stage investments in areas such as biotechnology, medical technology, AI, battery technology, semiconductors, and aerospace.[57]

USTC, one of the universities CAS manages, also funds a state-owned asset management company called USTC Holdings Company Limited (中科大资产经营有限责任公司), which is responsible for the commercialization of scientific research and management of university assets. According to its website, USTC Holdings owns stakes in more than 20 companies, including iFLYTEK.[58]

Commercialization: CAS University Spin-Off

Case Study: iFLYTEK

In 1999, a group of USTC students founded iFLYTEK—an AI company known for voice-recognition technology—using a prototype of speech synthesis technology first created at a lab at USTC.[59] Since the company's founding, iFLYTEK and USTC have maintained close research, workforce, and financial ties. For example, iFLYTEK and USTC jointly built the National Engineering Research Center of Speech and Language Information Processing (语音及语言信息处理国家工程研究中心), and in 2021 together participated in the Open Automatic Speech Recognition Challenge organized by the U.S. National Institute of Standards and Technology.[60] The two entities also maintain workforce- and education-development programs through on-campus training mechanisms and encouraging iFLYTEK employees to serve as part-time doctoral advisers.[61] USTC Holdings is iFLYTEK's fourth largest shareholder, and the chairman and founder of iFLYTEK holds professorial and supervisory positions at USTC and is the chairperson of CAS's Artificial Intelligence Industry-University-Research Innovation Alliance (中科院人工智能产学研创新联盟).[62] Notably, iFLYTEK is listed on the U.S. Commerce Department's Bureau of Industry and Security's (BIS) Entity List for its role in enabling the surveillance of Uyghurs in Xinjiang.[63]

Some CAS institutes also manage their own investment companies. For example, the Institute of Computing Technology—which is on the BIS Entity List "for acquiring and attempting to acquire U.S.-origin items in support of China's military modernization"—operates Beijing Zhongke Suanyuan Asset Management Co., Ltd. (北京中科算源资产管理有限公司), a wholly owned asset management company focused on computing technology investments.[64] According to its website, the company has helped the Institute of Computing Technology launch a number of China's computing and microelectronics companies, including supercomputing manufacturer Sugon (曙光信息产业股份有限公司), AI chip developer Cambricon (中科寒武纪科技股份有限公司), and CPU designer Loongson (龙芯中科技术股份有限公司), all of which are on the BIS Entity List.[65] Connections to CAS institutes afford these spin-off companies access to technical expertise, scientific facilities, increased resources, and other intangible benefits. In turn, successful spin-off companies often invest in and support other promising companies.[66]

Commercialization: CAS Institute Spin-Off

Case Study: Lenovo

In 1984, the CAS Institute of Computing Technology provided seed funding to 11 institute researchers to found the New Technology Development Company (中国科学院计算所新技术发展公司). In 1988, NTD Co. was reorganized into Legend Computer Group Co., which later rebranded to Lenovo Group (联想集团), or Lenovo in 2003.[67] The company's first product was an extension card, which is a hardware component that enhances a computer's functionality. This innovation allowed the company to secure financial backing, partner with firms to establish a manufacturing joint venture, and expand into Hong Kong.[68] In 1994, Legend Holdings, owner of the company, transferred 65 percent of its shares to CAS.[69] CAS's backing helped the company expand its operations, enter international markets, and eventually go public.[70] By 2004, Lenovo had become the world's third-largest PC company.[71] Today, Lenovo is one of the world's largest consumer electronics manufacturers.[72] Legend Holdings maintains a 31 percent stake in Lenovo Group; in turn, CAS maintains a 29 percent share of Legend Holdings.[73]

The case studies of iFLYTEK and Lenovo help reveal CAS's role in fostering the commercialization of critical technologies. CAS promotes the development of technology companies by providing financial support and personnel to help found them. Once established, these companies often maintain strong connections with CAS through joint research projects, financial ties, and talent development programs. Still, while these examples provide important insights into how CAS has contributed to China's technology commercialization, more work is needed to better understand CAS's overall contributions to facilitating tech transfer in China.

Advising S&T Policymaking

Beyond advancing S&T research and facilitating technological commercialization, CAS is a powerful player in China's S&T policymaking process and has a long history of shaping some of the country's most important scientific initiatives. For example, CAS recommendations resulted in the creation of the 863 Program (also known as the National High-Tech Development Plan), which fueled progress in supercomputing and aerospace technologies.[74] A CAS proposal also led to the development of the 973 Program, which was a key source of funding for basic research in important S&T fields until it was folded into China's National Key R&D Program in 2016.[75] In addition to being a powerful stakeholder in China's S&T bureaucracy, CAS influences the Chinese S&T policymaking process through its academicians (院士) and think tanks.

CAS academicians are among the most influential individuals in China's S&T ecosystem. They affect the allocation of resources to various research and commercial units, manage significant budgets, and in recent years have featured prominently in China's foremost political offices.[76] For instance, Yin Hejun (阴和俊), an academician who was previously the vice president of CAS, currently leads the Ministry of Science and Technology.[77] A total of 20 CAS academicians are currently in the 205-member 20th Central Committee, up from 18 in the 19th Central Committee (2017–2022) and eight in the 18th Central Committee (2012–2017).[78] CAS academicians do not need to hold official government positions to shape S&T policy. For instance, recommendations authored by CAS academicians resulted in the founding of the NSFC and the Chinese Academy of Engineering, both major players in China's S&T ecosystem today.[79]

Academicians also lead the research conducted by major science and technology think tanks such as the Academic Divisions of the Chinese Academy of Sciences (CASAD; 中国科学院学部) and the Institutes of Science and Development (CASISD; 中国科学院科技战略咨询研究院).[80] Established in 1955, CASAD advises the State Council and other government agencies on the formation and coordination of S&T policy and is intended to increase communication between researchers and policymakers.[81] Starting in 2019, for instance, CASAD conducted joint research with the NSFC to study future development paths of various emerging technologies and scientific fields of research critical to China's development.[82] The Research on the Development Strategy of Chinese Disciplines and Frontier Fields (中国学科及前沿领域发展战略研究) project brought together over three thousand scientists, including four hundred of the roughly eight hundred total CAS academicians, resulting in the publication of 38 books, each of which outlines a development strategy for a scientific discipline or frontier technology field.[83]

CASISD, founded in 2016, emerged out of a Chinese government initiative to form twenty-five "New Think Tanks with Chinese Characteristics" (中国特色新型智库) and a directive from Xi Jinping for CAS to create a "first-class" S&T think tank.[84] It is intended to improve China's assessment and policymaking capabilities.[85] CASISD uses its ties to CAS to assess progress in China's S&T ecosystem and makes S&T policy recommendations for both CAS and China generally.[86] According to the CAS 13th Five-Year Plan, CASISD is also tasked with promoting international scientific exchanges.[87]

Conclusion

The top global producer of STEM papers and home to many of the world's leading researchers, CAS is one of the most important S&T institutions in the world and plays a pivotal role in advancing China's S&T development. Our analysis of CAS's research output underscores the institution's key position in China's STEM ecosystem. CAS also fosters the commercialization and adoption of important technologies. Moreover, CAS actors, including academicians and think tanks, often help shape Chinese S&T policy.

To be sure, CAS has made progress in achieving the organizational goals set out in various policy documents.[88] CAS is a leading producer of impactful research and is the world's largest producer of highly cited and top-cited STEM publications. Successful tech companies have emerged from CAS research and investments, contributing to China's competitiveness in critical and emerging technologies. National S&T policies are often developed with input from CAS policymakers and academicians.

At the same time, CAS continues to face various challenges. While the organization has rapidly climbed global research rankings in recent years, only a handful of researchers from CAS have been awarded the world's most prestigious scientific awards.[89] No CAS scientist has ever received the Turing Award or Nobel Prize, for example. In addition, ongoing reforms to China's S&T bureaucracy, first announced in March 2023, could affect CAS's core functions and responsibilities. Intended to centralize and streamline S&T decision-making authorities under the newly announced Central Science and Technology Commission, these reforms will likely impact CAS's role in shaping national S&T policy.[90]

Furthermore, CAS's core functions—advancing research, commercializing technologies, and shaping S&T policy—may be in tension with one another. As a basic research organization, CAS is incentivized to meet or exceed its research goals. As a commercial entity, CAS must promote the growth of its spin-off enterprises and prioritize applied research projects that might lead to greater revenues for the organization. As an influential policymaking actor, CAS's leaders are incentivized to increase the organization's relative political power and influence within China's S&T bureaucracy. If CAS over-prioritizes its commercial activities, CAS scientists might invest less time in basic research in favor of projects that can be more easily commercialized. On the other hand, if CAS over-invests in developing its basic R&D programs at the expense of its commercial endeavors, it may negatively affect CAS's business interests, which are an essential source of revenue for the organization. CAS is tasked with shaping Chinese S&T policies, but it is also invested in protecting its own organizational interests, and therefore it may advocate against policies that threaten to weaken its position in

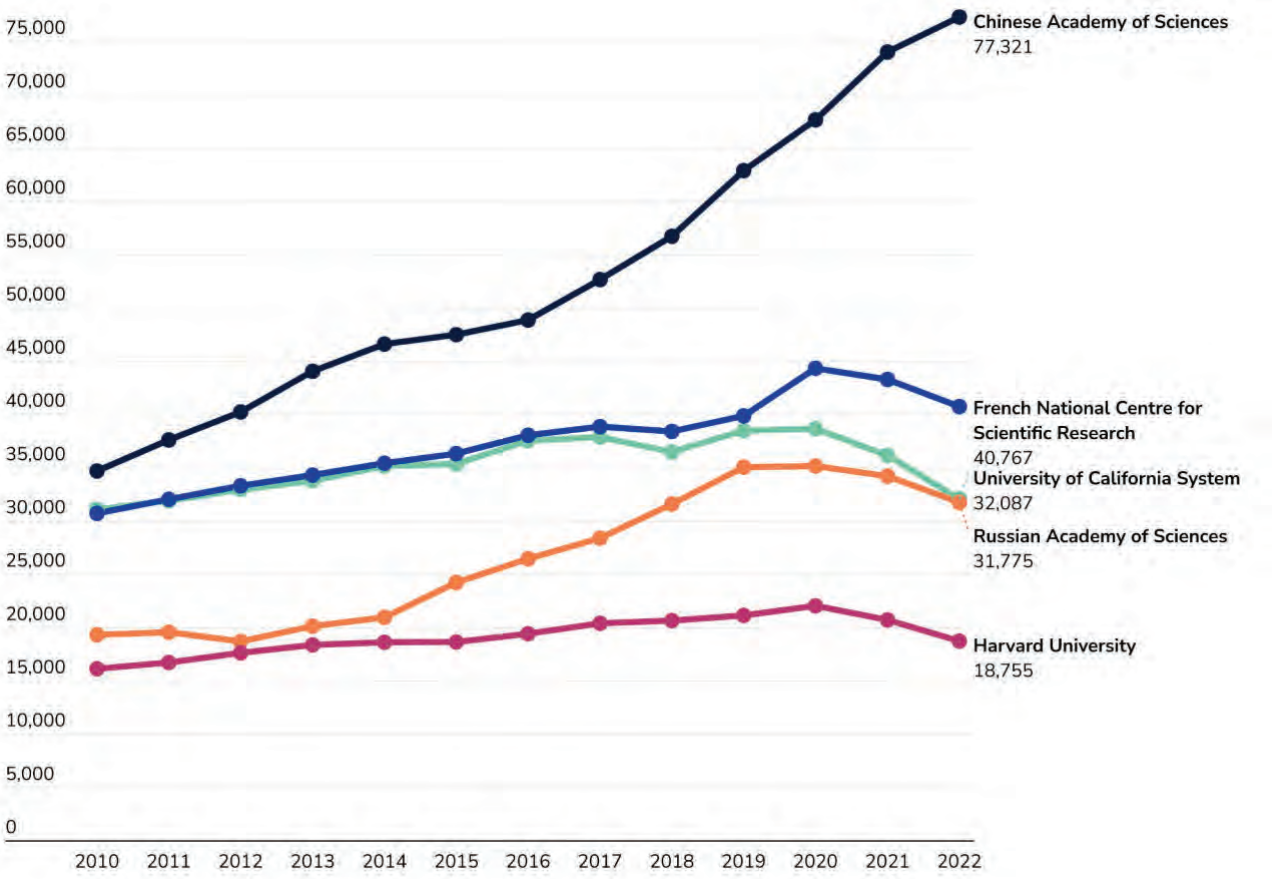
China's S&T ecosystem, even if those policies might better advance the country's S&T development.

The tensions among CAS's core functions described above could pose challenges to the organization's ability to simultaneously advance all three functions. Moreover, CAS's success in certain metrics—such as research output and number of top-cited researchers—relative to other leading STEM research institutions does not necessarily suggest that China's model of S&T development is more effective than that of other countries. As shown in this report, Beijing has committed extensive resources to CAS to push forward its S&T development goals. Consequently, CAS's success, or lack thereof, will affect Beijing's ability to achieve its own strategic objectives.

Despite these challenges, CAS's importance to Beijing's S&T ambitions underscores the need for U.S. policymakers to better understand the organization. While this report does not offer an exhaustive account of CAS's activities, it provides an important foundation for understanding CAS, which will remain a key driver of China's S&T development for the foreseeable future.

Appendix A: Additional Figures

Figure A1. Annual STEM Publications, Top Publishing Organizations, 2010–2022



Source: CSET Merged Corpus.

Figure A2. Top Publishing Organizations, Total and Highly Cited STEM Papers, 2010–2022

ALL STEM PAPERS

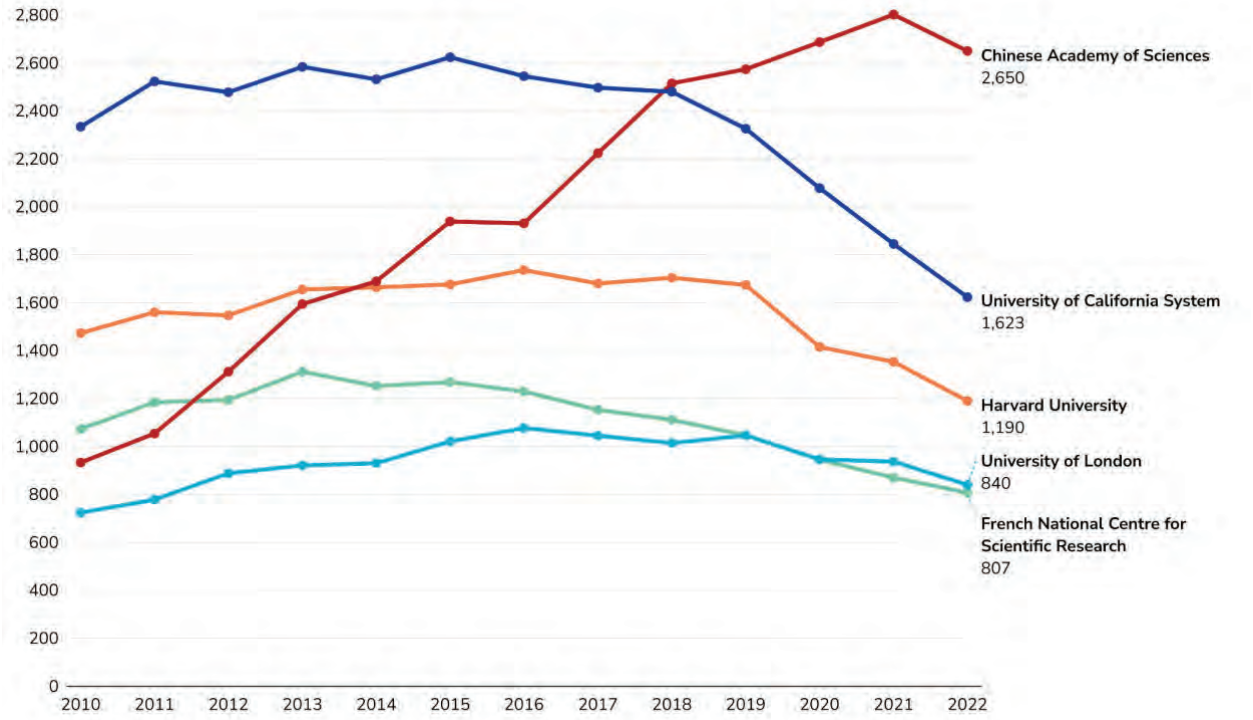
Rank	Organization	Papers
1	Chinese Academy of Sciences	759,271
2	French National Centre for Scientific Research	520,482
3	University of California System	479,045
4	Russian Academy of Sciences	366,441
5	Harvard University	265,000
6	United States Department of Energy	222,098
7	UDICE-French Research Universities	221,770
8	University of Texas System	213,960
9	University of London	212,462
10	Helmholtz Association	211,084
11	Indian Institutes of Technology System	194,378
12	Shanghai Jiao Tong University	187,910
13	Zhejiang University	187,475
14	Max Planck Society	181,472
15	Tsinghua University	175,969

HIGHLY CITED STEM PAPERS

Rank	Organization	Papers
1	Chinese Academy of Sciences	191,451
2	University of California System	169,959
3	French National Centre for Scientific Research	135,335
4	Harvard University	94,180
5	UDICE-French Research Universities	76,079
6	University of London	72,917
7	United States Department of Energy	72,243
8	The University of Texas System	66,782
9	Helmholtz Association	66,690
10	Max Planck Society	61,751
11	Massachusetts Institute of Technology	49,670
12	Stanford University	48,736
13	Spanish National Research Council	46,966
14	Pennsylvania Commonwealth System of Higher Education	46,895
15	Tsinghua University	46,356

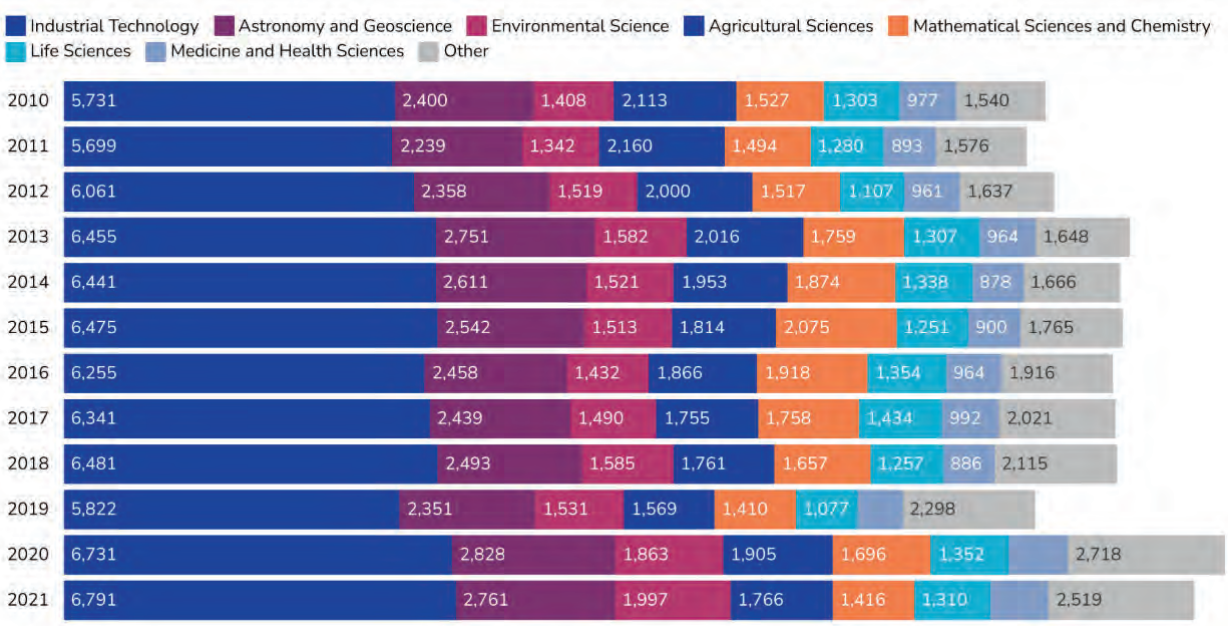
Source: CSET Merged Corpus.

Figure A3. Top Publishing Organizations by Top-Cited STEM Papers, 2010–2022



Source: CSET Merged Corpus.

Figure A4. CAS Institutes' Paper Publications by Fields of Study, 2010–2021



Source: CNKI.

Figure A5. AI-Related Papers by Top Publishers in China, STEM Papers, 2010–2021

	Organization	Papers
1	CAS Institutes	23,431
2	Wuhan University	9,520
3	University of Chinese Academy of Sciences	9,305
4	Nanjing University of Aeronautics and Astronautics	8,902
5	Shanghai Jiao Tong University	8,277
6	Tsinghua University	8,208
7	Northwestern Polytechnical University	8,206
8	Zhejiang University	7,974
9	Harbin Institute of Technology	7,626
10	Beihang University	6,925

Source: CNKI.

Appendix B: List of CAS Institutes

Name (Chinese)	Name (English)	Location	U.S. Government Restriction[91]
数学与系统科学研究院	Academy of Mathematics and Systems Science	Beijing	None
物理研究所	Institute of Physics	Beijing	Entity List[92]
理论物理研究所	Institute of Theoretical Physics	Beijing	None
高能物理研究所	Institute of High Energy Physics	Beijing	None
力学研究所	Institute of Mechanics	Beijing	None
声学研究所	Institute of Acoustics	Beijing	None
理化技术研究所	Technical Institute of Physics and Chemistry	Beijing	None
化学研究所	Institute of Chemistry	Beijing	None
国家纳米科学中心	National Center for Nanoscience and Technology	Beijing	None
生态环境研究中心	Research Center for Eco-Environmental Sciences	Beijing	None
过程工程研究所	Institute of Process Engineering	Beijing	None
地理科学与资源研究所	Institute of Geographic Sciences and Natural Resources Research	Beijing	None
国家天文台	National Astronomical Observatories	Beijing	None
云南天文台	Yunnan Observatories[93]	Kunming	None
南京天文光学技术研究所	Nanjing Institute of Astronomical Optics & Technology[94]	Nanjing	Entity List[95]
新疆天文台	Xinjiang Astronomical Observatory[96]	Urumqi	None
长春人造卫星观测站	Changchun Observatory[97]	Changchun	None

地质与地球物理研究所 (地球科学研究院）	Institute of Geology and Geophysics (Innovation Academy for Earth Science)	Beijing	None
青藏高原研究所	Institute of Tibetan Plateau Research	Lhasa	None
古脊椎动物与古人类研究所	Institute of Vertebrate Paleontology and Paleoanthropology	Beijing	None
大气物理研究所	Institute of Atmospheric Physics	Beijing	None
植物研究所	Institute of Botany	Beijing	None
动物研究所	Institute of Zoology	Beijing	None
心理研究所	Institute of Psychology	Beijing	None
微生物研究所	Institute of Microbiology	Beijing	None
生物物理研究所	Institute of Biophysics	Beijing	None
遗传与发育生物学研究所	Institute of Genetics and Developmental Biology (IGDB)[98]	Shijiazhuang	None
农业资源研究中心	Center for Agricultural Resources Research, (IGDB)[99]	Shijiazhuang	None
北京基因组研究所 (国家生物信息中心)	Beijing Institute of Genomics (China National Center for Bioinformation)	Beijing	None
计算技术研究所	Institute of Computing Technology	Beijing	Entity List[100]
软件研究所	Institute of Software	Beijing	None
半导体研究所	Institute of Semiconductors	Beijing	None
微电子研究所	Institute of Microelectronics[101]	Guangzhou	None
空天信息创新研究院	Aerospace Information Research Institute	Beijing	None
自动化研究所	Institute of Automation	Beijing	None
电工研究所	Institute of Electrical Engineering	Beijing	None
工程热物理研究所	Institute of Engineering Thermophysics	Beijing	None

国家空间科学中心	National Space Science Center	Beijing	None
自然科学史研究所	Institute for the History of Natural Sciences	Beijing	None
科技战略咨询研究院	Institutes of Science and Development	Beijing	None
信息工程研究所	Institute of Information Engineering	Beijing	None
数据与通信保护研究教育中心	Data Assurance & Communications Security Center[102]	Beijing	None
空间应用工程与技术中心	Technology and Engineering Center for Space Utilization	Beijing	None
天津工业生物技术研究所	Tianjin Institute of Industrial Biotechnology	Tianjin	None
大连化学物理研究所	Dalian Institute of Chemical Physics	Dalian	None
金属研究所	Institute of Metal Research	Shenyang	None
沈阳应用生态研究所	Institute of Applied Ecology	Shenyang	None
沈阳自动化研究所	Shenyang Institute of Automation	Shenyang	Entity List[103]
海洋研究所	Institute of Oceanology	Qingdao	None
青岛生物能源与过程研究所	Qingdao Institute of Bioenergy and Bioprocess Technology	Qingdao	None
烟台海岸带研究所	Yantai Institute of Coastal Zone Research	Yantai	None
长春光学精密机械与物理研究所	Changchun Institute of Optics, Fine Mechanics, and Physics	Changchun	None
长春应用化学研究所	Changchun Institute of Applied Chemistry	Changchun	None
东北地理与农业生态研究所	Northeast Institute of Geography and Agroecology	Changchun	None
农业技术中心	Center for Agricultural Technology[104]	Changchun	None

上海微系统与信息技术研究所	Shanghai Institute of Microsystem and Information Technology	Shanghai	Entity List[105]
上海技术物理研究所	Shanghai Institute of Technical Physics	Shanghai	None
上海光学精密机械研究所	Shanghai Institute of Optics and Fine Mechanics	Shanghai	None
上海硅酸盐研究所	Shanghai Institute of Ceramics	Shanghai	None
上海有机化学研究所	Shanghai Institute of Organic Chemistry	Shanghai	None
上海应用物理研究所	Shanghai Institute of Applied Physics	Shanghai	Unverified List[106]
上海天文台	Shanghai Astronomical Observatory	Shanghai	None
分子细胞科学卓越创新中心 (生物化学与细胞生物学研究所)	Center for Excellence in Molecular Cell Science (Shanghai Institute of Biochemistry and Cell Biology)	Shanghai	None
脑科学与智能技术卓越创新中心 (神经科学研究所)	Center for Excellence in Brain Science and Intelligence Technology (Institute of Neuroscience)	Shanghai	None
分子植物科学卓越创新中心 (植物生理生态研究所)	Center for Excellence in Molecular Plant Sciences (Institute of Plant Physiology and Ecology)	Shanghai	None
上海营养与健康研究所	Shanghai Institute of Nutrition and Health	Shanghai	None
上海药物研究所	Shanghai Institute of Materia Medica	Shanghai	None
上海免疫与感染研究所	Shanghai Institute of Immunity and Infection	Shanghai	None
上海高等研究院	Shanghai Advanced Research Institute	Shanghai	None

微小卫星创新研究院	Innovation Academy for Microsatellites	Shanghai	None
福建物质结构研究所 (海西研究院)	Fujian Institute of Research on the Structure of Matter (Haixi Institutes)	Fuzhou	None
宁波材料技术与工程研究所	Ningbo Institute of Materials Technology & Engineering	Ningbo	None
城市环境研究所	Institute of Urban Environment	Xiamen	None
杭州医学研究所	Hangzhou Institute of Medicine	Hangzhou	None
南京地质古生物研究所	Nanjing Institute of Geology and Palaeontology	Nanjing	None
南京土壤研究所	Institute of Soil Science	Nanjing	None
南京地理与湖泊研究所	Nanjing Institute of Geography & Limnology	Nanjing	None
紫金山天文台	Purple Mountain Observatory	Nanjing	None
苏州纳米技术与纳米仿生研究所	Suzhou Institute of Nano-tech and Nano-bionics	Suzhou	None
苏州生物医学工程技术研究所	Suzhou Institute of Biomedical Engineering and Technology	Suzhou	None
赣江创新研究院	Ganjiang Innovation Academy	Ganzhou	None
合肥物质科学研究院	Hefei Institutes of Physical Science	Hefei	None
武汉岩土力学研究所	Institute of Rock and Soil Mechanics	Wuhan	None
精密测量科学与技术创新研究院	Innovation Academy for Precision Measurement Science and Technology	Wuhan	None
武汉病毒研究所 (生物安全大科学研究中心)	Wuhan Institute of Virology (Center for Biosafety Mega-science)	Wuhan	None
水生生物研究所	Institute of Hydrobiology	Wuhan	None

武汉植物园	Wuhan Botanical Garden	Wuhan	None
南海海洋研究所	South China Sea Institute of Oceanology	Guangzhou	None
华南植物园	South China Botanical Garden	Guangzhou	None
广州能源研究所	Guangzhou Institute of Energy Conversion	Guangzhou	None
广州地球化学研究所	Guangzhou Institute of Geochemistry	Guangzhou	None
长沙矿产资源勘查中心	Changsha Center for Mineral Resource Exploration[107]	Changsha	None
广州生物医药与健康研究院	Guangzhou Institutes of Biomedicine and Health	Guangzhou	None
深圳先进技术研究院	Shenzhen Institute of Advanced Technology	Shenzhen	None
亚热带农业生态研究所	Institute of Subtropical Agriculture	Changsha	None
深海科学与工程研究所	Institute of Deep-Sea Science and Engineering	Wuhan	None
成都生物研究所	Chengdu Institute of Biology	Chengdu	None
成都山地灾害与环境研究所	Institute of Mountain Hazards and Environment	Chengdu	None
光电技术研究所	Institute of Optics and Electronics	Chengdu	None
重庆绿色智能技术研究院	Chongqing Institute of Green and Intelligent Technology	Chongqing	None
昆明动物研究所	Kunming Institute of Zoology	Kunming	None
昆明植物研究所	Kunming Institute of Botany	Kunming	None
西双版纳热带植物园	Xishuangbanna Tropical Botanical Garden	Xishuangbanna	None
广州地球化学研究所	Guangzhou Institute of Geochemistry	Guangzhou	None

西安光学精密机械研究所	Xi'an Institute of Optics and Precision Mechanics	Xi'an	None
国家授时中心	National Time Service Center	Xi'an	None
地球环境研究所	Institute of Earth Environment	Xi'an	None
山西煤炭化学研究所	Institute of Coal Chemistry	Taiyuan	None
近代物理研究所	Institute of Modern Physics	Lanzhou	None
兰州化学物理研究所	Lanzhou Institute of Chemical Physics	Lanzhou	None
西北生态环境资源研究院	Northwest Institute of Eco-environment and Resources	Lanzhou	None
青海盐湖研究所	Qinghai Institute of Salt Lakes	Xining	None
西北高原生物研究所 (三江源国家公园研究院)	Northwest Institute of Plateau Biology (Institute of Sanjiangyuan National Park)	Xining	None
新疆理化技术研究所	Xinjiang Technical Institute of Physics and Chemistry	Urumqi	None
新疆生态与地理研究所	Xinjiang Institute of Ecology and Geography	Urumqi	None

Source: Chinese Academy of Sciences.

Authors

Cole McFaul is a research analyst at CSET and a nonresident fellow at the Atlantic Council's Global China Hub.

Hanna Dohmen is a research analyst at CSET, a nonresident fellow at the Atlantic Council's Global China Hub, and consults for Covington & Burling LLP on semiconductor policy issues.

Sam Bresnick is a research fellow at CSET.

Emily Weinstein contributed to this report while she was a research fellow at CSET. She is currently detailed to the U.S. Department of Commerce under an Intergovernmental Personnel Act agreement. Her contributions to this report were completed prior to her service in the Commerce Department. The views expressed herein are those of the authors and do not necessarily reflect those of the U.S. government.

Acknowledgments

The authors would like to thank Emmy Probasco, Margarita Konaev, Catherine Aiken, Daniel Chou, Ben Murphy, and Ngor Luong for their feedback and assistance. The authors are also grateful for the reviews and suggestions they received from John Chen, Cong Cao, Dakota Cary, Bill Hannas, and Huey-Meei Chang. They also thank Shelton Fitch and Liz Dana for editorial support, and Jason Ly for help with graphic design.

This report was updated on October 25, 2024, to clarify the relationship between CAS, Legend Holdings, and Lenovo.

Document Identifier: doi: 10.51593/20220055

Endnotes

[1] The State Council (国务院), also known as the Central People's Government of the People's Republic of China, is the top executive organ of the Chinese government. See: William Hannas, James Mulvenon, and Anna Puglisi, *Chinese Industrial Espionage: Technology Acquisition and Military Modernisation* (New York: Routledge, 2013).

[2] "2023 Tables: Institutions," Nature Index, www.nature.com/nature-index/annual-tables/2023/institution/all/all/global.

[3] For more, see: Autumn Toney and James Dunham, "Multi-label Classification of Scientific Research Documents Across Domains and Languages," *Association for Computational Linguistics Proceedings of the Third Workshop on Scholarly Document Processing* (October 2022): 105–114.

[4] For more on how we classify AI-related research, see: Daniel Chou, "Counting AI Research" (CSET, July 2022), https://cset.georgetown.edu/wp-content/uploads/CSET-Counting-AI-Research.pdf.

[5] Hannas, Mulvenon, and Puglisi, *Chinese Industrial Espionage*.

[6] Paul H. B. Godwin, "Strategic Forces," in *Development of the Chinese Armed Forces* (Montgomery, AL: Air University Press, 1988), https://irp.fas.org/dia/product/prc_ch-11.htm.

[7] Godwin, "Strategic Forces."

[8] Liu Yanqiong, "Chinese Academy of Sciences and the 'Two Bombs and One Satellite' Project," *Proceedings of the Chinese Academy of Sciences* 34, no. 9 (2019): 1003–1013. The Chinese term here, 两弹一星, literally means "two bombs and one satellite." It refers to China's Mao Zedong–era accomplishments of building its own nuclear bombs (原子弹), ballistic missiles (导弹), and earth satellites (卫星). The "two bombs" (两弹) part is often misconstrued as "atomic and hydrogen bombs." See William Hannas and Huey-Meei Chang, "China's STI Operations: Monitoring Foreign Science and Technology Through Open Sources" (CSET, January 2021), 27, https://cset.georgetown.edu/publication/chinas-sti-operations.

[9] "Computer 156 [156 计算机]," *CCF China Computer Historical Memory*, 2018, https://perma.cc/2Y3Q-VWQ3.

[10] The six institutes were the Institute of Computing Technology, Institute of Electronics, Institute of Physics, Northeast Institute of Physics, Northwest Institute of Computing Technology, and Institute of Applied Chemistry. Today, as part of the China Aerospace Science and Technology Corporation (CASC), the Xi'an Microelectronics Technology Institute remains a key People's Liberation Army supplier and manages two military-standard integrated circuit foundries. CASC is a state-owned defense and aerospace conglomerate. For more, see Eric Lee and Seamus Boyle, "Chinese Nuclear Missile Guidance Systems: Spotlight on the Xian Institute of Microelectronics Technology" (Project 2049 Institute, September 18, 2020), https://project2049.net/2020/09/18/chinese-nuclear-missile-guidance-systems-spotlight-on-the-xian-institute-of-microelectronics-technology/.

[11] Xi'an Microelectronics Technology Institute still holds 34 percent ownership of ZTE, and financial filings show that ZTE maintains close ties to the institute. The institute is the second largest shareholder of ZTE. The company and institute are also tied through shared personnel. The institute in 2022 nominated three of the nine directors on the company's board, and the current chairman and executive director of ZTE previously held numerous leadership positions at the institute, including deputy head of from 2015 to 2019. In August 2022, the U.S. Department of Commerce's Bureau of Industry and Security (BIS) added Xi'an Microelectronics Technology Institute to the Entity List "for acquiring and attempting to acquire US-origin items in support of China's military modernization efforts." Lee and Boyle, "Chinese Nuclear Missile Guidance Systems"; "2022 Annual Report" (ZTE Corporation, March 2023), https://perma.cc/7ZN8-X42W; BIS, "Additions of Entities to the Entity List," *Federal Register* 87 FR 51876 (August 24, 2022), www.federalregister.gov/documents/2022/08/24/2022-18268/additions-of-entities-to-the-entity-list.

[12] Richard P. Suttmeier, Cong Cao, and Denis Fred Simon, "China's Innovation Challenge and the Remaking of the Chinese Academy of Sciences," *Innovations: Technology, Governance, Globalization* 1, no. 3 (Summer 2006): 78–97, https://doi.org/10.1162/itgg.2006.1.3.78.

[13] Suttmeier, Cao, and Simon, "China's Innovation Challenge."

[14] Suttmeier, Cao, and Simon, "China's Innovation Challenge."

[15] Richard P. Suttmeier, Cong Cao, and Denis Fred Simon, "'Knowledge Innovation' and the Chinese Academy of Sciences," *Science* 312, no. 5770 (April 7, 2006): 58–59.

[16] Suttmeier, Cao, and Simon, "Knowledge Innovation." For example, in 2001, the Institute of Genetics and Developmental Biology was formed by merging the Institute of Genetics and the Institute of Developmental Biology. In 2003, the Beijing Genomics Institute—now BGI Group—was separated from the Institute of Genetics and Developmental Biology and established as a CAS research institute. "Institute Overview" [研究所概况], Institute of Genetics and Developmental Biology, accessed May 9, 2024, https://perma.cc/PMC3-SPX6.

[17] Suttmeier, Cao, and Simon, "Knowledge Innovation"; Federal Bureau of Investigation Strategic Partnership Unit, *Chinese Talent Programs*, Counterintelligence Strategic Partnership Intelligence Note (Washington, DC: FBI, September 2015), https://info.publicintelligence.net/FBI-ChineseTalentPrograms.pdf.

[18] The seven new institutes established were the Guangzhou Institutes of Biomedicine and Health, Institute of Urban Environment (in Xiamen), Yantai Institute of Coastal Zone Research, Suzhou Institute of Nano-tech and Nano-bionics, Qingdao Institute of Bioenergy and Bioprocess Technology, Ningbo Institute of Materials Technology & Engineering, and the Shenzhen Institute of Advanced Technology. Micah Springut, Stephen Schlaikjer, and David Chen, *China's Program for Science and Technology Modernization: Implications for American Competitiveness* (Washington, DC: U.S.-China Economic and Security Review Commission, April 20, 2011), www.uscc.gov/sites/default/files/Research/USCC_REPORT_China%27s_Program_forScience_and_Technology_Modernization.pdf.

[19] Springut, Schlaikjer, and Chen, *China's Program for Science and Technology Modernization*.

[20] Jane Qiu, "China Sets 2020 Vision for Science," *Science* 470, no. 15 (February 1, 2011), https://doi.org/10.1038/470015a.

[21] Yuehui Wu, "Chinese Academy of Sciences Contributes to Belt and Road Construction through Science Cooperation," *People's Daily*, April 23, 2019, https://perma.cc/7LYW-J93M.

[22] CAS, "Outline of the Chinese Academy of Sciences 13th Five-Year Development Plan" [中国科学院'十三五'发展规划纲要], trans. CSET (CSET, October 20, 2022), https://cset.georgetown.edu/publication/outline-of-the-chinese-academy-of-sciences-13th-five-year-development-plan/.

[23] Jane Qiu, "Chinese Academy of Sciences Has Big Plans for Nation's Research," *Nature*, March 24, 2011, www.nature.com/articles/news.2011.180.

[24] CAS, "Chinese Academy of Sciences 2022 Budget" [中国科学院 2022 年部门预算], trans. CSET (CSET, June 30, 2022), https://cset.georgetown.edu/publication/chinese-academy-of-sciences-2022-budget/.

[25] USTC (中国科学技术大学) was founded in 1958 to support China's "Nuclear Bombs, Ballistic Missiles, and Earth Satellites" program. USTC houses 32 colleges focusing on cutting-edge science and technology, medicine, and humanities subjects. USTC also has five research institutes: the Suzhou Institute for Advanced Research, the Shanghai Institute for Advanced Studies, Beijing Research Institute, the Institute of Advanced Technology, and the International Institute of Finance. On May 9, 2024, USTC was added to the BIS's Entity List for acquiring and attempting to acquire U.S.-origin items in support of advancing China's quantum technology capabilities and for being involved in advancing China's nuclear program development.

UCAS (中国科学院大学)—as it is known today—was established in 2012, but the school has existed since 1978. Formerly known as the Graduate School of the University of Science and Technology of China, it was the first graduate school in China. Since 2014, however, UCAS has enrolled both undergraduate and graduate students.

ShanghaiTech University (上海科技大学) was founded in 2013 as a collaboration between the Shanghai Municipal People's Government and CAS. The university consists of five schools and three research institutes. ShanghaiTech claims to be an internationally focused university that aims to cultivate innovative scientists, investors, and entrepreneurs. The majority of classes are taught in both Chinese and English, and the university heavily emphasizes study abroad opportunities at prestigious universities, including Harvard University, the Massachusetts Institute of Technology, and Yale University. According to its website, 14 percent of faculty are foreign nationals and almost 25 percent of its graduates pursue postgraduate programs overseas.

"School Profile" [学校简介], USTC, January 2024, https://perma.cc/E643-UF2N; BIS, "Additions of Entities to the Entity List," *Federal Register* 89 FR 41886 (May 09, 2024), www.federalregister.gov/public-inspection/2024-10485/additions-of-entities-to-the-entity-list; "School

Profile" [学校简介], UCAS, accessed April 10, 2024, https://perma.cc/P8BX-Q6ZV; "School Profile" [学校简介], ShanghaiTech University, accessed April 10, 2024, https://perma.cc/QVD4-FJTL.

[26] Emily Weinstein, Channing Lee, Ryan Fedasiuk, and Anna Puglisi, "China's State Key Laboratory System: A View into China's Innovation System" (CSET, June 2022), https://doi.org/10.51593/20210019.

[27] Weinstein et al., "China's State Key Laboratory System."

[28] "About Us," CAS, accessed May 9, 2024, https://perma.cc/D3MV-GLFX; "Publications," CAS, accessed May 9, 2024, https://perma.cc/3K5Q-2DKD; "Machine Intelligence Research," Springer Link, accessed May 9, 2024, https://link.springer.com/journal/11633; "Journal of Plant Ecology," Oxford University Press, accessed May 9, 2024, https://academic.oup.com/jpe.

[29] For example, as of 2020, USTC had signed cooperation agreements with 39 CAS institutes and more than 30 CAS academicians and institute directors serve as part-time department leaders at the university. See "Introduction to 'All of CAS Helps Run the School and Its Departments Are Combined with CAS Institutes'" ['全院办校、所系结合'简介], USTC, August 2020, https://archive.ph/xYYBd.

[30] "Introduction to the Chinese Academy of Sciences" [中国科学院简介], CAS, August 2022, https://perma.cc/UJ86-KP9Q; "Who We Are," CAS, accessed April 10, 2024, https://perma.cc/CQV5-NZK6.

[31] The NSFC is a government-run body that oversees and audits Chinese scientific research funds, principally those that support basic research. CAS, "About Us"; "Overview" [概况], NSFC, accessed August 19, 2024, https://perma.cc/LGG9-QD2T.

[32] CAS, "The Chinese Academy of Sciences (CAS)," International Research Collaboration Information Platform, May 27, 2016, https://perma.cc/3RSV-K9U6; David Cyranoski, "China Fires Up Next-Generation Neutron-Science Facility," *Nature* 551, no. 284 (November 16, 2017), https://doi.org/10.1038/nature.2017.22976; "Another World Record for China's EAST Tokamak," *Nuclear Engineering International*, April 18, 2023, www.neimagazine.com/news/newsanother-world-record-for-chinas-east-tokamak-10768385; Institute of Plasma Physics, "Experimental Advanced Superconducting Tokamak (EAST)," CAS Large Research Infrastructures User Service Platform, accessed April 16, 2024, https://perma.cc/55SU-NANM.

[33] USD conversions used average exchange rates for years 2021 and 2023. Currency exchange rates retrieved from "Chinese Yuan to US Dollar Spot Exchange Rates for 2021," Exchange Rates UK, accessed April 10, 2024, www.exchangerates.org.uk/CNY-USD-spot-exchange-rates-history-2021.html; "Chinese Yuan to US Dollar Spot Exchange Rates for 2023," Exchange Rates UK, accessed April 10, 2024, www.exchangerates.org.uk/CNY-USD-spot-exchange-rates-history-2023.html.

[34] Figure 1 was created using data collected from the CAS budgets, which are published annually, and national basic R&D expenditure data published by the National Bureau of Statistics of China. Chinese organizations refer to the *Frascati Manual* of the Organization for Economic Cooperation and Development to define basic R&D activities. National Bureau of Statistics of China, "Notice on Issuing the 'Standards for the Statistics of Research and Experimental Development (R&D) Investment (Trial)'"

[国家统计局关于印发《研究与试验发展（R&D）投入统计规范（试行）》的通知], *National Bureau of Statistics*, no. 47 (April 19, 2019), https://perma.cc/9J4S-Q76H.

[35] As of 2023, according to the National Bureau of Statistics of China, Chinese basic R&D expenditure had risen to 6.6 percent of total R&D spending. "Cross-National Comparisons of R&D Performance," in *Science & Engineering Indicators 2018* (Washington, DC: National Science Board, January 2018), www.nsf.gov/statistics/2018/nsb20181/report/sections/research-and-development-u-s-trends-and-international-comparisons/cross-national-comparisons-of-r-d-performance.

[36] "U.S. R&D Expenditures," in *National Patterns of R&D Resources 2021–2022* (Arlington, VA: National Center for Science and Engineering Statistics, January 2024), https://ncses.nsf.gov/data-collections/national-patterns/2021-2022#data. Both the United States and China use OECD's definitions for basic R&D activities. "Wan Donghua, the Main Responsible Person of the Department of Social, Science, Culture and Health Statistics of the National Bureau of Statistics, Answers Reporters' Questions Regarding the Release of the 'Standards for Statistics on Research and Experimental Development (R&D) Input (Trial)'" [国家统计局社科文司主要负责人万东华就发布《研究与试验发展（R&D）投入统计规范（试行）》答记者问], Xinjiang Survey Corps of the National Bureau of Statistics, May 20, 2019, https://perma.cc/BB24-9384; OECD, "Concepts and Definitions for Identifying R&D," in *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development* (Paris: OECD Publishing, October 8, 2015), 47–70, https://doi.org/10.1787/9789264239012-4-en.

[37] Nature Index, "2023 Tables: Institutions."

[38] Nature Index, "2023 Tables: Institutions."

[39] See, for example, Dag W. Aksnes, Liv Langfeldt, and Paul Wouters, "Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories," *Sage Open* 9, no. 1 (2019), https://doi.org/10.1177/2158244019829575.

[40] Ken Hyland, "Enter the Dragon: China and Global Academic Publishing," *Learned Publishing* 36, no. 3 (May 18, 2023), https://doi.org/10.1002/leap.1545.

[41] Hyland, "Enter the Dragon."

[42] "Highly Cited Researchers 2023," Clarivate, accessed April 10, 2024, https://clarivate.com/highly-cited-researchers/.

[43] Clarivate, "Highly Cited Researchers 2023."

[44] Anderson, "U.S. R&D Increased by $72 Billion in 2021 to $789 Billion"; "Chinese Academy of Sciences 2022 Department Budget" [中国科学院 2022 年部门预算] (Beijing: CAS), https://perma.cc/UFQ5-KWJP.

[45] National Center for Science and Engineering Statistics, "Harvard U: Headcount of R&D Personnel: 2022–16," National Science Foundation, accessed August 19, 2024, https://ncsesdata.nsf.gov/profiles/site?method=report&tin=U1300001&id=h4.

[46] University of California, "The University of California at a Glance," March 2024, https://ucop.edu/institutional-research-academic-planning/_files/uc-facts-at-a-glance.pdf.

[47] CAS, "Outline of the Chinese Academy of Sciences 13th Five-Year Development Plan."

[48] Fang Xu and Xiaoxuan Li, "The Changing Role of Metrics in Research Institute Evaluations Undertaken by the Chinese Academy of Sciences (CAS)," *Palgrave Communications*, no. 2 (October 25, 2016), https://doi.org/10.1057/palcomms.2016.78.

[49] "Yang Chunli" [杨春莉], UCAS, accessed June 18, 2024, https://perma.cc/K4XL-9GYU .

[50] "Introduction to University of Chinese Academy of Sciences (UCAS)," UCAS, accessed June 18, 2024, https://perma.cc/2SZY-CGGN.

[51] See, for example, Institute of Automation, "Admission Brochure for the Institute of Automation, Chinese Academy of Sciences, to Recruit Master's Degree Students in 2024" [中国科学院自动化研究所 2024 年招收攻读硕士学位研究生招生简章], CAS, September 22, 2023, https://perma.cc/YE2N-R9CW.

[52] Liu Gang et al., *China's New Generation Artificial Intelligence Technology Industry Development Report 2022* [中国新一代人工智能科技产业发展报告 2022] (Chinese Institute of New Generation Artificial Intelligence Development Strategies [中国新一代人工智能发展战略研究院], June 24, 2022), https://perma.cc/3259-XU3J.

[53] CAS, "Outline of the Chinese Academy of Sciences 13th Five-Year Development Plan." "Hidden Champion" (隐形冠军) enterprises prioritize R&D and focus on developing cutting-edge products in niche markets. Source: https://perma.cc/T4MH-4CN2.

[54] Cong Cao and Richard P. Suttmeier, "Challenges of S&T System Reform in China," *Science* 355, no. 6329 (April 10, 2017): 1019–2021, https://doi.org/10.1126/science.aal2515; "Notice of the State Council on Issuing and Implementing Certain Provisions of the 'Law of the People's Republic of China on Promoting the Transformation of Scientific and Technological Achievements'" [国务院关于印发实施《中华人民共和国促进科技成果转化法》若干规定的通知], State Council, no. 16 (March 2, 2016), https://cset.georgetown.edu/publication/state-council-notice-on-the-publication-of-certain-regulations-on-implementing-the-law-of-the-peoples-republic-of-china-on-promoting-the-conversion-of-scientific-and-technological-achievements/.

[55] Suttmeier, Cao, and Simon, "'Knowledge Innovation"; "Notice of the Chinese Academy of Sciences on the Issuance of the 'Measures for the Openness and Sharing of Large-Scale Scientific Research Instruments of the Chinese Academy of Sciences'" [中国科学院关于印发《中国科学院大型科研仪器开放共享管理办法》的通知], CAS, March 25, 2022, https://perma.cc/J2GC-MQZS; State Council, "Opinions of the State Council on the Opening of Major National Scientific Research Infrastructure and Large-Scale Scientific Research Instruments to the Public" [国务院关于国家重大科研基础设施和大型科研仪器向社会开放的意见], Science and Technology Innovation and Development Center, December 31, 2014, https://perma.cc/TUW2-85XR; 科技部, 发展改革委, and 财政部, "Notice of the Development and Reform Commission of the Ministry of Science and Technology and the Ministry of Finance on the Issuance of

the 'Measures for the Openness and Sharing of National Major Scientific Research Infrastructure and Large Scientific Research Instruments" [科技部 发展改革委 财政部关于印发《国家重大科研基础设施和大型科研仪器开放共享管理办法》的通知], 中国科学院条件保障与财务局, September 20, 2017, 国科发基〔2017〕289 号 edition, https://perma.cc/S87C-C88W.

[56] "Corporate Bond Annual Report (2022)" [公司债券年度报告 (2022 年)] (Chinese Academy of Sciences Holdings Co., Ltd., April 2023), https://perma.cc/EG53-K8WD.

[57] "Who Are We?" [我们是谁?], CAS Investment Management Co., Ltd, accessed April 16, 2024, https://perma.cc/3MV3-SQ2B.

[58] "Company Introduction" [公司简介], USTC Holdings Company Limited, accessed April 15, 2024, https://perma.cc/R32R-GVEM.

[59] "We Are Very Proud That iFLYTEK Originated from the University of Science and Technology of China" [我们很自豪，讯飞源于中科大!], National Engineering Research Center of Speech and Language Information Processing, February 28, 2022, https://perma.cc/3L6X-2XQT.

[60] "2021 Annual Report" (iFLYTEK Co., Ltd., 2021), https://perma.cc/RD9E-PEQT.

[61] National Engineering Research Center of Speech and Language Information Processing, "We Are Very Proud."

[62] iFLYTEK, "2021 Annual Report."

[63] BIS, "Additions of Entities to the Entity List," *Federal Register* 87 FR 51876 (October 9, 2019), www.federalregister.gov/documents/2022/08/24/2022-18268/additions-of-entities-to-the-entity-list.

[64] BIS, "Additions and Revisions to the Entity List and Conforming Removal from the Unverified List," *Federal Register* 87 FR 77505 (December 16, 2022), www.federalregister.gov/documents/2022/12/19/2022-27151/additions-and-revisions-to-the-entity-list-and-conforming-removal-from-the-unverified-list.

[65] "Business Incubation" [企业孵化], Institute of Computing Technology, accessed April 15, 2024, https://perma.cc/UZS4-VCL7.

[66] For example, Lenovo Venture Capital was a seed investor in chip designer Jeejio (中科物栖) in 2018 and Sugon is a 50 percent shareholder in data intelligence firm Golaxy (中科天玑). "Jeejio" [中科物栖], 36Krypton Venture Capital Platform, accessed April 16, 2024, https://perma.cc/7928-EWLM; "Golaxy" [中科天玑], 36Krypton Venture Capital Platform, accessed April 16, 2024, https://perma.cc/QYL3-EKQP.

[67] "History" [历史沿革], Institute of Computing Technology, accessed April 15, 2024, https://perma.cc/3MSH-86F4; "Development Path" [发展历程], Legend Holdings, accessed June 18, 2024, https://perma.cc/GT56-KXRD.

[68] Nathaniel Ahrens and Yu Zhou, "China's Competitiveness: Myths, Realities, and Lessons for the United States and Japan—Case Study: Lenovo" (Center Strategic & International Studies, January 2013), https://www.csis.org/programs/japan-chair/japan-chair-archives/chinas-competitiveness-myths-realities-and-lessons-united.

[69] While we have not found unambiguous documentation regarding the origins of Legend Holdings, we do know that in 1994 Legend Holdings was the owner of the company producing "Legend" branded computers and in that year transferred 65 percent of its own shares to CAS. Ahrens and Zhou, "China's Competitiveness."

[70]Ahrens and Zhou, "China's Competitiveness"; Institute of Computing Technology, "History" [历史沿革].

[71] Ahrens and Zhou, "China's Competitiveness."

[72] "Lenovo - Statistics and Facts," Statista, accessed October 24, 2024, https://perma.cc/YWL2-FLAR.

[73] "2023 Annual Report," (Legend Holdings), https://perma.cc/3UVU-6RBE; "Lenovo Group Limited 2023/24 Annual Report," (Lenovo Group), https://perma.cc/T9LS-493G.

[74] Xiaoxuan Li, Kejia Yang, and Xiaoxi Xiao, "Scientific Advice in China: The Changing Role of the Chinese Academy of Sciences," *Nature*, July 12, 2016, www.nature.com/articles/palcomms201645; CAS, "About Us"; "About CASAD," Academic Divisions of the Chinese Academy of Sciences, accessed April 15, 2024, https://archive.ph/8Ze5l.

[75] CAS, "About Us."

[76] Jane Qiu, "Chinese Academies Promise Cleaner Elections," *Nature*, August 10, 2011, www.nature.com/articles/476139a; Hao Xin, "The True Cost of Becoming an Academician in China?," *Science*, September 17, 2013, www.science.org/content/article/true-cost-becoming-academician-china.

[77] "Yin Hejun" [阴和俊], Ministry of Science and Technology, accessed April 15, 2024, https://perma.cc/P3AQ-UXUT; Jiang Chenglong, "Yin Hejun Appointed as Minister of Science and Technology," *China Daily*, October 24, 2023, https://perma.cc/UP3C-A2SG.

[78] CSET analysis of the Central Committee; Xinhua, "List of Members of 20th CPC Central Committee," State Council Information Office, October 22, 2022, https://perma.cc/XSV2-3LTU.

[79] The Chinese Academy of Engineering, often referred to with CAS as the "two academies," is a parallel organization to CAS focused on engineering sciences. Although smaller than CAS, CAE's structure and functions are similar, except focused in engineering sciences. "About Us," CAE, accessed June 17, 2024, https://en.cae.cn/cae/html/en/col2014/column_2014_1.html.

[80] The CASAD claims it is comparable to the U.S. National Academy of Sciences. CASAD, "About CASAD."

[81] Alex Stone, "China's Model of Science: Rationale, Players, Issues" (China Aerospace Studies Institute, 2022), www.airuniversity.af.edu/Portals/10/CASI/documents/Research/Infrastructure/2022-02-07%20Model%20of%20Science.pdf.

[82] "An Overview of China's 2035 Development Strategy for Academic Disciplines and Frontier Fields" [中国学科及前沿领域 2035 发展战略总论] (NSFC and CAS, March 2023), https://perma.cc/T9L8-C3KS.

[83] "'2035 Development Strategy Series of Chinese Disciplines and Frontier Fields' Published" ["中国学科及前沿领域 2035 发展战略丛书"出版] (CASAD, May, 26, 2023, https://perma.cc/753W-4VDV; "An Overview of China's 2035 Development Strategy."

[84] "National High-End Think Tank Construction Pilot Project" [国家高端智库建设试点], China Development Institute, accessed April 15, 2024, https://perma.cc/JA9N-NUXU; "Review and Research on the First Anniversary of the National High-End Think Tank Construction Pilot Work" [国家高端智库建设试点工作一周年回顾与研究], *Chinese Communist Party News Network, People's Daily*, December 1, 2016, https://perma.cc/K3JS-HRJY.

[85] China Development Institute, "National High-End Think Tank Construction Pilot Project."

[86] Science and Technology Daily, "The First Phase of the 'Take the Lead Initiative' Plan Has Yielded Fruitful Results" [【科技日报】"率先行动"计划第一阶段硕果累累], CAS, accessed April 15, 2024, https://archive.ph/VcXb9; "Introduction to the Institute of Science and Technology Strategy Consulting, Chinese Academy of Sciences" [中国科学院科技战略咨询研究院简介], Institutes of Science and Development, accessed April 15, 2024, https://perma.cc/9EBR-VMJT.

[87] CAS, "Outline of the Chinese Academy of Sciences 13th Five-Year Development Plan."

[88] Every five years, CAS publishes a development plan that sets the organization's goals for the period. See, for example, CAS, "Outline of the Chinese Academy of Sciences 13th Five-Year Development Plan."

[89] Lokman Meho identified one hundred of the world's most prestigious scientific awards. Just 11 of the 3,445 awards in Meho's dataset featured a researcher with a CAS affiliation. Lokman I. Meho, "Highly Prestigious International Academic Awards and Their Impact on University Rankings," *Quantitative Science Studies* 1, no. 2 (2020): 824–848, https://doi.org/10.1162/qss_a_00045. Data accessible from: Lokman I. Meho, "Highly Prestigious International Academic Awards and Their Impact on University Rankings (Article and Supporting Data)," American University in Beirut, February 12, 2020, https://scholarworks.aub.edu.lb/handle/10938/21535.

[90] Barry Naughton, Tai Ming Cheung, Siwen Xiao, Yaosheng Xu, and Yujing Yang, "Reorganization of China's Science and Technology System" (UC Institute on Global Conflict and Cooperation, July 2023), https://ucigcc.org/wp-content/uploads/2023/08/Naughton-et-al-Working-Paper-Reorg-v1-8.22.23.pdf.

[91] We used the International Trade Administration's Consolidated Screening List to determine whether a CAS institute faces U.S. government restrictions. This tool checks against BIS's Denied Persons List, Unverified List, Entity List, and Military End User List; the Department of State Bureau of International Security and Nonproliferation's Nonproliferation Sanctions; the Department of State Directorate of

Defense Trade Controls' Arms Export Control Act Debarred List; and the Department of Treasury Office of Foreign Assets Control's Specially Designated Nationals (SDN) List, Foreign Sanctions Evaders List, Sectoral Sanctions Identification List, Correspondent Account or Payable-Through Account Sanctions List, Non-SDN Menu-Based Sanctions List, Non-SDN Chinese Military-Industrial Complex Companies, and Palestinian Legislative Council List.

[92] The Institute of Physics was added to the Entity List on May 9, 2024, for acquiring and attempting to acquire U.S.-origin items in support of advancing China's quantum technology capabilities. BIS, "Additions of Entities to the Entity List," *Federal Register* 89 FR 41886 (May 9, 2024), www.federalregister.gov/public-inspection/2024-10485/additions-of-entities-to-the-entity-list.

[93] The Yunnan Observatories are affiliates of the Beijing-based National Astronomical Observatories (国家天文台), but they are located in Kunming, Yunnan. "Introduction" [单位简介], National Astronomical Observatories, accessed May 9, 2024, https://archive.ph/JEOpB.

[94] The Nanjing Institute of Astronomical Optics and Technology is an affiliate of the Beijing-based National Astronomical Observatories (国家天文台), but it is located in Nanjing, Jiangsu. National Astronomical Observatories, "Introduction" [单位简介].

[95] The Nanjing Institute of Astronomical Optics & Technology was added to the Entity List on September 27, 2023, for procuring U.S.-origin items in likely furtherance of Chinese military research. BIS, "Addition of Entities and Revision to Existing Entities on the Entity List; Removal of Existing Entity from the Military End User List," *Federal Register* 88 FR 66271 (September 27, 2023), www.federalregister.gov/documents/2023/09/27/2023-21080/addition-of-entities-and-revision-to-existing-entities-on-the-entity-list-removal-of-existing-entity.

[96] The Xinjiang Astronomical Observatory is an affiliate of the Beijing-based National Astronomical Observatories (国家天文台), but it is located in Urumqi, Xinjiang. National Astronomical Observatories, "Introduction" [单位简介].

[97] The Changchun Observatory (literally, "Changchun Artificial Satellite Observation Station") is an affiliate of the Beijing-based National Astronomical Observatories (国家天文台), but it is located in Changchun, Jilin. National Astronomical Observatories, "Introduction" [单位简介].

[98] The Beijing Genomics Institute—now BGI Group—was part of the Institute of Genetics and Developmental Biology until 2003, when it was separated into a separate institute. Institute of Genetics and Developmental Biology, "Institute Overview" [研究所概况].

[99] The Center for Agriculture Resources Research is nominally a component of the Beijing-based CAS Institute of Genetics and Developmental Biology (遗传与发育生物学研究所), but it is located in a different city (Shijiazhuang, Hebei) and is legally a separate entity. "Introduction to the Center" [中心简介], Center for Agricultural Resources Research, accessed May 7, 2024, https://perma.cc/LR9C-LKVW.

[100] The Institute of Computing Technology was added to the Entity List on December 16, 2022, for acquiring and attempting to acquire U.S.-origin items in support of China's military modernization. BIS identifies this entity as a major AI-chip R&D, manufacturing, and sales entity. They also identify this

entity for being, or having close ties to, government organizations that support the Chinese military and the defense industry. BIS, "Additions and Revisions to the Entity List and Conforming Removal from the Unverified List," *Federal Register* 87 FR 77505 (December 16, 2022), www.federalregister.gov/documents/2022/12/19/2022-27151/additions-and-revisions-to-the-entity-list-and-conforming-removal-from-the-unverified-list.

[101] The predecessor of the Institute of Microelectronics was the CAS 109 Factory (中科院 109 工厂). In 1958, it established China's first transistor production factory. In 1965, the CAS 109 Factory built the 109B computer, which contributed to the development of the hydrogen bomb. In 1986, the CAS 109 Factory merged with the Institute of Semiconductors and the Institute of Computing Technology to form the Institute of Microelectronics. Liu "Chinese Academy of Sciences and the 'Two Bombs and One Satellite' Project"; "Institute Overview" [研究所概况], Institute of Microelectronics, accessed May 9, 2024, https://perma.cc/9A52-7EX8.

[102] The Data Assurance & Communications Security Center is one part of the State Key Laboratory of Information Security (信息安全国家重点实验室). Both the DCS Center and its parent laboratory are components of the CAS Institute of Information Engineering (中国科学院信息工程研究所). "Introduction to the DCS Center" [DCS 中心介绍], Data Assurance & Communications Security Center, accessed May 7, 2024, https://perma.cc/JV6V-LGD6.

[103] The Shenyang Institute of Automation was added to the Entity List on June 8, 2022, for acquiring and/or attempting to acquire U.S.-origin items in support of military applications, contrary to the national security or foreign policy interests of the United States. BIS, "Addition of Entities, Revision and Correction of Entries, and Removal of Entities from the Entity List," *Federal Register* 87 FR 38920 (June 28, 2022), www.federalregister.gov/documents/2022/06/30/2022-14069/addition-of-entities-revision-and-correction-of-entries-and-removal-of-entities-from-the-entity-list.

[104] The Center for Agricultural Technology is a component of the CAS Northeast Institute of Geography and Agroecology (中国科学院东北地理与农业生态研究所). "Homepage" [首页], Center for Agricultural Technology, accessed May 7, 2024, https://perma.cc/NGA6-XWMK.

[105] The Shanghai Institute of Microsystem and Information Technology was added to the Entity List on May 9, 2024, for acquiring and attempting to acquire U.S.-origin items in support of advancing China's quantum technology capabilities. BIS, "Additions of Entities to the Entity List," *Federal Register* 89 FR 41886 (May 9, 2024), www.federalregister.gov/public-inspection/2024-10485/additions-of-entities-to-the-entity-list.

[106] The Shanghai Institute of Applied Physics was added to the Unverified List on April 11, 2019, because BIS was unable to verify their bona fides through an end-use check. BIS, "Revisions to the Unverified List (UVL)," *Federal Register* 84 FR 14608 (April 11, 2019), www.federalregister.gov/documents/2019/04/11/2019-07211/revisions-to-the-unverified-list-uvl.

[107] The Changsha Center for Mineral Resource Exploration is a component of the CAS Guangzhou Institute of Geochemistry (广州地球化学研究所). "Changsha Center for Mineral Resource Exploration" [长沙矿产资源中心], Guangzhou Institute of Geochemistry, accessed May 9, 2024, https://perma.cc/ABD8-JMF5.

Exhibit 27

Article

Effect of non-pharmaceutical interventions to contain COVID-19 in China

https://doi.org/10.1038/s41586-020-2293-x

Received: 12 March 2020

Accepted: 23 April 2020

Published online: 4 May 2020

Check for updates

Shengjie Lai[1,2,10]✉, Nick W. Ruktanonchai[1,3,10]✉, Liangcai Zhou[4], Olivia Prosper[5], Wei Luo[6,7], Jessica R. Floyd[1], Amy Wesolowski[8], Mauricio Santillana[6,7], Chi Zhang[9], Xiangjun Du[9], Hongjie Yu[2] & Andrew J. Tatem[1]✉

On 11 March 2020, the World Health Organization (WHO) declared coronavirus disease 2019 (COVID-19) a pandemic[1]. The strategies based on non-pharmaceutical interventions that were used to contain the outbreak in China appear to be effective[2], but quantitative research is still needed to assess the efficacy of non-pharmaceutical interventions and their timings[3]. Here, using epidemiological data on COVID-19 and anonymized data on human movement[4,5], we develop a modelling framework that uses daily travel networks to simulate different outbreak and intervention scenarios across China. We estimate that there were a total of 114,325 cases of COVID-19 (interquartile range 76,776–164,576) in mainland China as of 29 February 2020. Without non-pharmaceutical interventions, we predict that the number of cases would have been 67-fold higher (interquartile range 44–94-fold) by 29 February 2020, and we find that the effectiveness of different interventions varied. We estimate that early detection and isolation of cases prevented more infections than did travel restrictions and contact reductions, but that a combination of non-pharmaceutical interventions achieved the strongest and most rapid effect. According to our model, the lifting of travel restrictions from 17 February 2020 does not lead to an increase in cases across China if social distancing interventions can be maintained, even at a limited level of an on average 25% reduction in contact between individuals that continues until late April. These findings improve our understanding of the effects of non-pharmaceutical interventions on COVID-19, and will inform response efforts across the world.

As of 30 March 2020, the outbreak of COVID-19, which is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has resulted in 693,282 confirmed cases and 33,106 deaths across the world[6]. As the disease has only recently emerged, effective pharmaceutical interventions are not expected to be available for months[7], and healthcare resources will be limited for treating all cases. Non-pharmaceutical interventions (NPIs) are therefore essential components of the public health response to COVID-19 outbreaks[6,8–10]. These include the isolation of individuals who are ill, contact tracing, quarantine of exposed individuals, travel restrictions, school and workplace closures, cancellation of mass gatherings, and hand-washing, among others[8–10]. Such measures aim to reduce the transmission of the virus by delaying the timing and reducing the size of the peak of the epidemic, thus buying time for preparations to be made in the healthcare system and creating the potential for vaccines and drugs to be used at a later stage[8].

Three major groups of NPIs have been implemented to contain the spread and reduce the size of the outbreak of COVID-19 across China[11]. First, intercity travel restrictions were used to prevent further seeding of the virus during the Chinese New Year holiday period. A cordon sanitaire of Wuhan and surrounding cities in Hubei province was put in place on 23 January 2020, two days before the Chinese New Year, which started on 25 January 2020. After this date, travel restrictions were also put in place in other provinces across the country. Second, the early identification and isolation of cases was prioritized, including improving the screening, identification, diagnosis, isolation, reporting and contact tracing of people who were suspected or confirmed to have the disease[11]. Local governments across China encouraged and supported the routine screening and quarantine of travellers from Hubei province in an attempt to detect COVID-19 infections as early as possible. The average interval from the onset of symptoms to laboratory confirmation dropped from 12 days in the early stages of the outbreak to 3 days in early February, indicating that these efforts improved detection and diagnosis[3,12]. Third, contact restrictions and social distancing measures, together with personal preventive actions such as hand-washing, were implemented to reduce the risk of exposure at the community level. As part of these social distancing policies, the Chinese government encouraged people to stay at home as much as possible, cancelled or postponed large public events and mass gatherings, and closed libraries, museums and workplaces[13,14]. School holidays were also extended, with the end date of the Chinese New Year holiday period changed from 30 January 2020 to 10 March 2020 for Hubei province, and to 9 February 2020 for many other provinces[15,16].

[1]WorldPop, School of Geography and Environmental Science, University of Southampton, Southampton, UK. [2]School of Public Health, Fudan University, Key Laboratory of Public Health Safety, Ministry of Education, Shanghai, China. [3]Population Health Sciences, Virginia Tech, Blacksburg, VA, USA. [4]Wuhan Center for Disease Control and Prevention, Wuhan, China. [5]Department of Mathematics, University of Tennessee, Knoxville, TN, USA. [6]Computational Health Informatics Program, Boston Children's Hospital, Boston, MA, USA. [7]Department of Pediatrics, Harvard Medical School, Boston, MA, USA. [8]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. [9]School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, China. [10]These authors contributed equally: Shengjie Lai, Nick W. Ruktanonchai. ✉e-mail: shengjie.lai@soton.ac.uk; nr1e14@soton.ac.uk; a.j.tatem@soton.ac.uk

The implementation of these NPIs coincided with a rapid decline in the number of new cases across China, albeit at high economic and social costs[3,12]. Previous studies have examined the effects of the lockdown of Wuhan[17,18], travel restrictions[19], airport screening[20], isolation of cases and contact tracing on the containment of the disease[21]. However, a comprehensive and quantitative comparison of the effectiveness of different NPIs, and the time at which they were implemented, for containing the outbreak of COVID-19 in China is lacking. On the basis of epidemiological data on COVID-19 and historical and near-real-time anonymized data on human movement, we developed a stochastic susceptible–exposed–infectious–removed (SEIR) modelling framework based on travel networks to simulate the spread of COVID-19 across 340 prefecture-level cities in mainland China. Within each city, we estimated the numbers of susceptible, exposed, infectious, and recovered/removed ('removed' refers to the individuals who were isolated to prevent further transmission, and deceased individuals) people per day from 1 December 2019. Using this modelling framework, we conducted before-and-after comparable analyses to quantify the relative effect of the three major groups of NPIs—that is, the restriction of intercity population movement, the identification and isolation of cases, and the reduction of travel and contact within cities to increase social distance—in China. We also assessed the risk of COVID-19 transmission since the lifting of travel restrictions on 17 February 2020.

Reconstructing the spread of COVID-19

The epidemiological parameters that were estimated for the early stage of the outbreak in Wuhan were initially used to parameterize the epidemic before interventions were widely implemented[5]. The three major groups of NPIs outlined above were derived and measured using data on population movement between and within cities (obtained from smartphone users of Baidu location-based services[4]) and data on the delay between the onset of illness and the reporting of cases across the country. Population travel and contact patterns changed substantially after the implementation of interventions, and the timeliness of case reporting also improved (Fig. 1, Supplementary Tables 1, 2). These indicators were then incorporated into the model (see Methods).

We estimated that there were a total of 114,325 cases of COVID-19 (interquartile range (IQR) 76,776–164,576) in mainland China as of 29 February 2020, 85% of which were in Hubei province (Extended Data Table 1). The outbreak increased exponentially before Chinese New Year, but the peaks of epidemics across the country quickly appeared around the time of Chinese New Year after the implementation of NPIs. The estimated epidemics and peaks were consistent with patterns of reported data by onset date, with strong correlations between daily estimates and reported data across time and regions (Extended Data Fig. 1). The overall correlation between the estimated number of cases and the reported number by province, as of 29 February 2020, was also significant ($P < 0.001$, $R^2 = 0.86$), with a high sensitivity (91%, 280/308) and specificity (69%, 22/32) in predicting cities with or without cases of COVID-19 (Extended Data Fig. 1a, b).

Quantifying the effect of different NPIs

Without NPIs, our model predicted the number of cases of COVID-19 to increase rapidly across China, with a 51-fold (IQR 33–71) increase in Wuhan, a 92-fold (58–133) increase in other cities in Hubei province and a 125-fold (77–180) increase in other provinces by 29 February 2020. However, the apparent effectiveness of different interventions varied (Fig. 2). The lockdown of Wuhan might not have prevented the seeding of the virus from the city, as the travel ban was put in place at the latter stages of population movement out of the city before Chinese New Year[22] (Fig. 1b). Nevertheless, if intercity travel restrictions had not been implemented, cities and provinces outside of Wuhan would have received more cases from Wuhan, and the affected geographical range would have expanded to the remote western areas of China (Extended Data Fig. 2c). In general, we estimated that the early detection and isolation of cases quickly and substantially prevented more infections than did the introduction of contact reduction and social distancing measures across the country (5-fold versus 2.6-fold). However, without the contact reduction intervention, in the longer term the epidemics would have increased exponentially across regions (Fig. 2c, f). Therefore, combined NPIs would bring about the strongest and most rapid effect on containment of the COVID-19 outbreak, with an interval of about one week between the introduction of NPIs and the peak of the epidemic (Extended Data Table 1).

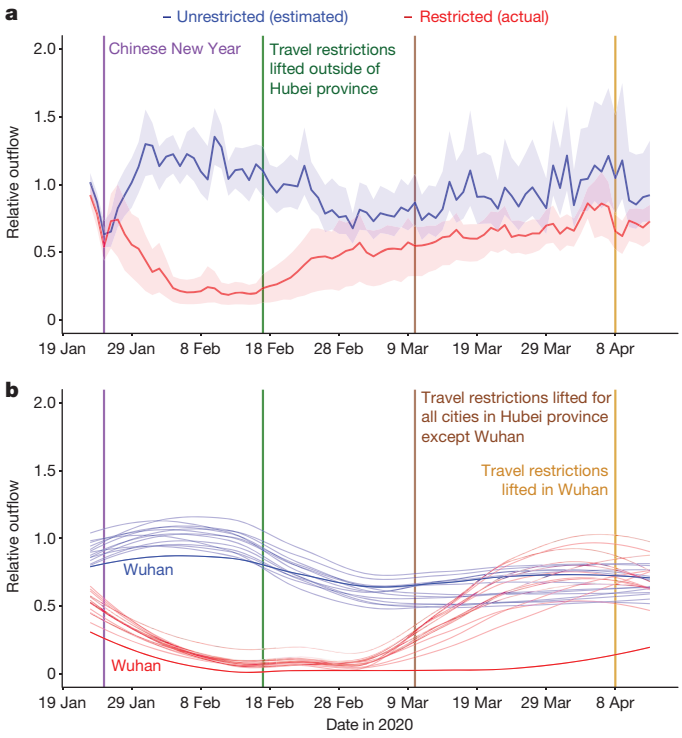


Fig. 1 | Relative daily volume of outbound travellers from cities across mainland China between 23 January and 13 April 2020. a, Relative outbound flows of travellers for all cities at prefecture level ($n = 340$) in mainland China, presented as the median (solid line) and IQR (shading). **b**, Relative outbound flows of travellers for cities in Hubei province. Wuhan is highlighted using darker colours. Each red line represents the outflow for each city in 2020, standardized by the mean of daily outflows for each city from 20 to 22 January 2020. Each blue line represents an estimate of the normal outflow by city under the scenario of no travel restrictions (on the basis of travel in previous years). The lines in **b** were smoothed using locally estimated scatterplot smoothing (LOESS) regression.

Timing of interventions

Our model suggests that, theoretically, if interventions in China had been implemented one week, two weeks or three weeks earlier than they actually were, the number of cases of COVID-19 could have been reduced by 66% (IQR 50–82%), 86% (81–90%) or 95% (93–97%), respectively (Fig. 3a). The geographical range of affected areas would also shrink from 308 cities to 192, 130 or 61 cities, respectively (Extended Data Fig. 3). However, if NPIs had been introduced one week, two weeks or three weeks later than they were, the number of cases might have increased by 3-fold (IQR 2–4), 7-fold (5–10) or 18-fold (11–26), respectively (Fig. 3b).

Lifting of travel restrictions

Under the interventions that were implemented from 17 February 2020—that is, the lifting of travel restrictions—the epidemics outside

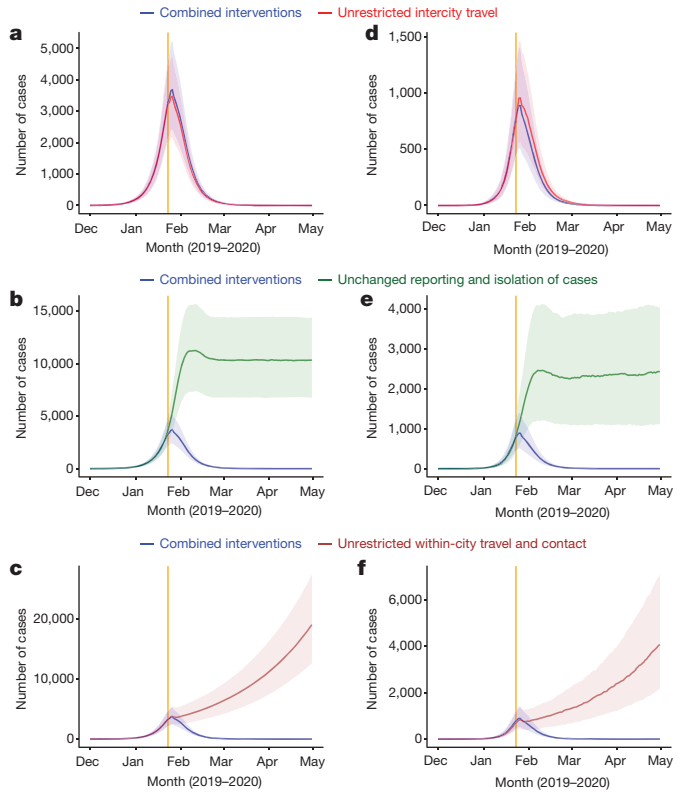


Fig. 2 | Estimated epidemic curves of the COVID-19 outbreak under various scenarios with or without NPIs by region. a–**c**, Estimates for the city of Wuhan. **d**–**f**, Estimates for cities outside of Hubei province in mainland China. The blue lines represent estimated transmission under combined NPIs, and the other coloured lines represent the scenario without one type of intervention. Data are presented as the median (solid line) and IQR (shading) of estimates (1,000 simulations). The orange vertical lines indicate the date on which the lockdown of Wuhan began (23 January 2020).

of Hubei province probably reached a low level (fewer than 10 cases per day, excluding imported cases from other countries) in early March, whereas Hubei province might need another four weeks to reach the same level as other provinces. However, if population contact resumed to normal levels, the lifting of travel restrictions might cause case numbers to rise again (Fig. 3c). Accordingly, our simulations suggest that maintaining social distancing even to a limited extent (for example, a 25% reduction in contact between individuals on average) through to late April would help to ensure control of COVID-19 in epicentres such as Wuhan.

Our estimates were sensitive to the basic reproduction number (R_0); under a higher R_0 value, the peaks of epidemics were higher and later, and more time was needed to contain the outbreak (Extended Data Fig. 3). Sensitivity analyses also suggested that our model could have robustly measured relative changes in the efficacy of interventions under different epidemiological parameters and transmission scenarios (Extended Data Figs. 4–9).

Discussion

Our findings show that combined NPIs substantially reduced the transmission of COVID-19 across China. Earlier implementation of NPIs could have notably reduced the magnitude and geographical range of the outbreak, but—equally—a delayed response would have led to a larger outbreak. China's aggressive, multifaceted response is likely to have prevented a far worse situation, which would have accelerated the spread of the virus globally. The evidence from China provides information that will be of use in efforts to contain the spread of COVID-19 and mitigate the effects of the disease in other regions around the world[3,12].

Our results suggest three key points. First, they support and validate the idea that population movement and close contact has a major role in the spread of COVID-19 within and beyond China[22,23]. As the lockdown of Wuhan happened at the latter stages of population movement before Chinese New Year, travel restrictions did not halt the seeding of the virus from Wuhan, but did prevent cases being exported from Wuhan to a wider area. Second, the importance and effects of the three types of NPIs differed. Compared with travel restrictions, improved detection and isolation of cases, as well as social distancing, probably had a greater effect on the containment of the outbreak. The social distancing intervention reduced contact between people who travelled from the epicentre of the outbreak and other individuals. This is likely to have been especially helpful in curbing the spread of an emerging pathogen to the wider community, and to have reduced the risk of spread from asymptomatic or mild infections[8]. Third, given that travel and work have begun to resume in China, the country should consider at least the partial continuation of NPIs to ensure that the COVID-19 outbreak is sustainably controlled for the first wave of this outbreak. For example, the early identification and isolation of cases should be maintained—which might also help to prevent and delay the arrival of a second wave, considering the increasing numbers of cases that are imported from other countries and the presence of asymptomatic or subclinical infections in China[24].

The analyses presented here provide a comprehensive quantitative assessment of the effect of NPIs on the transmission of COVID-19. The model framework accounts for daily interactions of populations and interventions between and within cities, as well as the inherent

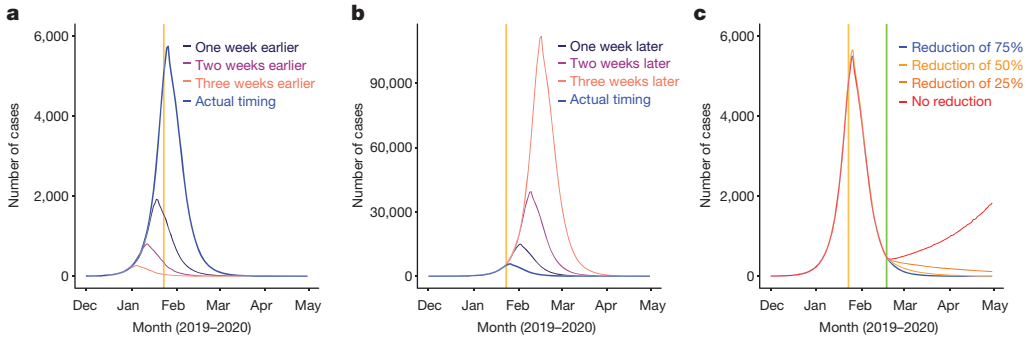


Fig. 3 | Estimates of the COVID-19 outbreak under various scenarios of intervention timing and lifting of travel restrictions across China. a, Estimated epidemic curves for interventions implemented earlier than their actual timing. **b**, Estimated epidemic curves for interventions implemented later than their actual timing. **c**, Estimated spread of COVID-19 for different rates of population contact after the lifting of intercity travel restrictions. The orange vertical lines indicate the date on which the lockdown of Wuhan began (23 January 2020), and the green line shows the date on which travel restrictions were lifted (17 February 2020).

statistical uncertainty that is associated with a paucity of epidemiological parameters before and after the implementation of interventions. The network-based SEIR model is methodologically robust and is built on the basic SEIR models that have been used previously to predict the transmission of COVID-19 in its early stages[23]. Considering the delays that exist in the reporting of cases, our approach can be used to enable a rapid, ongoing estimation of the effectiveness of various NPIs in different countries, and to aid decision-making relating to the control of outbreaks of COVID-19.

Our study has several limitations. First, our simulations were based on parameters that were estimated for symptomatic cases identified in the early stage of the outbreak in Wuhan, and might not account for asymptomatic and mild infections; we may therefore have underestimated the total number of infections. Second, our findings could be confounded by other factors that changed during the outbreak. Although we have shown that the apparent fall in the incidence of COVID-19 after Chinese New Year (25 January 2020) in China is likely to be attributed to the interventions taken, we cannot rule out the possibility that the decrease was partially attributable to other unknown seasonal factors—for example, temperature and absolute humidity[25,26]. Third, if the epidemiological parameters of COVID-19 transmission in other cities across China differed from the estimates[5]—which were based on the data in the early stage of the outbreak, when no NPIs were in place in Wuhan—then our estimates of the effectiveness of interventions in reducing the transmission of COVID-19 could be biased. Fourth, there are probably biases in population coverage, given that our model relies on data from mobile phone and Baidu users. Although a high percentage (from 46.9% in 2013 to 55.3% in 2018) of the population of China owns smartphones[27] (https://en.wikipedia.org/wiki/List_of_countries_by_smartphone_penetration), the group of mobile-phone users does not include specific subgroups of the population, particularly children. Therefore, our data on population movement may provide an incomplete picture, and differences between the characteristics of smartphone owners and non-owners may also bias our estimates. In addition, the magnitude and patterns of population movements could change year by year—although previous studies have suggested that travel patterns are consistent in their seasonality across years in China and other countries[22]. Finally, we only examined three main groups of NPIs, and other interventions might also have contributed to the containment of the outbreak. For example, owing to the sources of data that were available, we did not assess the effect of personal hygiene and protective equipment on containing the spread of COVID-19. Other sources of data and further investigations are needed to measure and evaluate the efficacy of each intervention.

COVID-19 has placed a substantial burden on health systems and society across many countries. From a public health standpoint, our results highlight that countries should consider proactively planning NPIs and relevant strategies for containment and mitigation, as the earlier implementation of NPIs could have led to substantial reductions in the size of the outbreak in China. Our results also provide guidance for countries as to the likely effectiveness of different NPIs at different stages of an outbreak. Suspected and confirmed cases of the disease should be identified, diagnosed, isolated and reported as early as possible to control the source of infection, and the implementation of cordon sanitaires or travel restrictions for areas that are heavily affected might prevent the virus spreading to wider regions. Reducing contact and increasing social distance between individuals, together with improved personal hygiene, can help to protect vulnerable populations and mitigate the spread of COVID-19 at the community level, and these interventions should be promoted throughout the outbreak to avoid resurgence. Our findings suggest that—as advocated by WHO—strategies that involve the early implementation of integrated NPIs should be prepared, deployed and adjusted to maximize the benefits of these interventions and minimize the health, social and economic effects of COVID-19 around the world[3].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2293-x.

1. World Health Organization. *WHO Director-General's opening remarks at the media briefing on COVID-19 – 11 March 2020* https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 (WHO, 2020).
2. Epidemiology Working Group for NCIP Epidemic Response, Chinese Center for Disease Control and Prevention. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China [in Chinese with English abstract]. *Zhonghua Liu Xing Bing Xue Za Zhi* **41**, 145–151 (2020).
3. World Health Organization. *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)* https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf (WHO, 2020).
4. *Baidu Migration* [in Chinese] https://qianxi.baidu.com/ (2020).
5. Li, Q. et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020).
6. World Health Organization. *Coronavirus disease (COVID-19) pandemic* https://www.who.int/emergencies/diseases/novel-coronavirus-2019 (WHO, 2020).
7. Heymann, D. L. & Shindo, N. COVID-19: what is next for public health? *Lancet* **395**, 542–545 (2020).
8. Fong, M. W. et al. Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings—social distancing measures. *Emerg. Infect. Dis.* **26**, 976–984 (2020).
9. Ryu, S. et al. Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings—international travel-related measures. *Emerg. Infect. Dis.* **26**, 961–966 (2020).
10. Xiao, J. et al. Nonpharmaceutical measures for pandemic influenza in nonhealthcare settings—personal protective and environmental measures. *Emerg. Infect. Dis.* **26**, 967–975 (2020).
11. Chen, W. et al. Early containment strategies and core measures for prevention and control of novel coronavirus pneumonia in China [in Chinese with English abstract]. *Zhonghua Yu Fang Yi Xue Za Zhi* **54**, 239–244 (2020).
12. World Health Organization. *Press Conference of WHO-China Joint Mission on COVID-19* https://www.who.int/docs/default-source/coronaviruse/transcripts/joint-mission-press-conference-script-english-final.pdf?sfvrsn=51c90b9e_2 (2020).
13. The State Council of the People's Republic of China. *The announcement from Wuhan's headquarter on the novel coronavirus prevention and control* [in Chinese] http://www.gov.cn/xinwen/2020–01/23/content_5471751.htm (2020).
14. The State Council of the People's Republic of China. *The announcement on strengthening community prevention and control of pneumonia epidemic situation of new coronavirus infection* [in Chinese] http://www.gov.cn/zhengce/2020-01/27/content_5472516.htm (2020).
15. The State Council of the People's Republic of China. *The State Council's announcement on extending the 2020 Spring Festival Holiday* [in Chinese] http://www.gov.cn/zhengce/content/2020-01/27/content_5472352.htm (2020).
16. The People's Government of Shanghai Municipality. *The announcement on postponing the reoperation date of companies and the reopening date of schools* [in Chinese] http://www.shanghai.gov.cn/nw2/nw2314/nw2315/nw43978/u21aw1423601.html (2020).
17. Li, X., Zhao, X. & Sun, Y. The lockdown of Hubei Province causing different transmission dynamics of the novel coronavirus (2019-nCoV) in Wuhan and Beijing. Preprint at *medRxiv* https://doi.org/10.1101/2020.02.09.20021477 (2020).
18. Kraemer, M. U. G. et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497 (2020).
19. Chinazzi, M. et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395–400 (2020).
20. Quilty, B. J., Clifford, S., CMMID nCoV working group 2, Flasche, S. & Eggo, R. M. Effectiveness of airport screening at detecting travellers infected with novel coronavirus (2019-nCoV). *Euro Surveill.* **25**, 2000080 (2020).
21. Hellewell, J. et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob. Health* **8**, e488–e496 (2020).
22. Lai, S. et al. Assessing spread risk of Wuhan novel coronavirus within and beyond China, January–April 2020: a travel network-based modelling study. Preprint at *medRxiv* https://doi.org/10.1101/2020.02.04.20020479 (2020).
23. Wu, J. T., Leung, K. & Leung, G. M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet* **395**, 689–697 (2020).
24. National Health Commission of the People's Republic of China. *Updates on pneumonia of new coronavirus infections as of March 31, 2020* [in Chinese] http://www.nhc.gov.cn/xcs/yqtb/202004/28668f987f3a4e58b1a2a75db60d8cf2.shtml (2020).
25. Wang, M. et al. Temperature significant change COVID-19 transmission in 429 cities. Preprint at *medRxiv* https://doi.org/10.1101/2020.02.22.20025791 (2020).
26. Luo, W. et al. The role of absolute humidity on transmission rates of the COVID-19 outbreak. Preprint at *medRxiv* https://doi.org/10.1101/2020.02.12.20022467 (2020).
27. Lai, S., Farnham, A., Ruktanonchai, N. W. & Tatem, A. J. Measuring mobility, disease connectivity and individual risk: a review of using mobile phone data and mHealth for travel medicine. *J. Travel Med.* **26**, taz019 (2019).

Methods

Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Model summary

An SEIR model based on travel networks was built to simulate the spread of COVID-19 between and within all prefecture-level cities in mainland China. This model has been made openly available for further use at https://github.com/wpgp/BEARmod. Population movement data across the country were used to estimate the intensity of travel restrictions and contact reductions. Data from illness onset to reporting of the first index case for each county were used to infer the changing timeliness of case identification and isolation across the course of the outbreak. The outputs of the model under NPIs were validated by using daily numbers of new cases reported across all regions in mainland China. On the basis of this modelling framework, the efficacy of applying or lifting non-pharmaceutical measures under various scenarios and timings were tested and quantified.

Data sources

Three datasets on population movement, which were obtained from Baidu location-based services that provide over 7 billion positioning requests per day[4,28], were used in this study to measure travel restrictions and social distancing across time and space. The first is an aggregated and de-identified dataset on near-real-time daily relative outbound and inbound flow of smartphone users for each prefecture-level city in 2020 (340 cities in mainland China were included) to understand patterns of mobility during the outbreak. The daily outflow from each city since the lockdown of Wuhan and the travel restrictions that were applied on 23 January 2020 were rescaled by the mean daily flow for each city from 20 to 22 January 2020 for comparing travel reductions across cities and years (Fig. 1).

The second Baidu dataset is a historical relative movement matrix with daily total number of users at the city level from 26 December 2014 to 26 May 2015, aligning with the 2020 Chinese New Year holiday period, for which the corresponding period is 1 December 2019 to 30 April 2020. We assumed that the pattern of population movements was the same in years when there were no outbreaks and interventions. Adjusted by the level of travel reductions derived from the 2020 dataset where applicable, the second dataset was used to simulate the spread of COVID-19 and predict transmission via population movements under various scenarios, with or without intercity travel restrictions. Corresponding city-level population data in 2015 for modelling were obtained from the Chinese Bureau of Statistics[29].

The third Baidu dataset measures daily population movements at the county level (2,862 counties in China) from 26 January to 30 April 2014, as described elsewhere[30]. On the basis of the assumption that the pattern of population contact was consistent across years when there were no interventions, it was used to estimate within-city travel and contact reduction during the outbreak and interventions. First, we aggregated data from county to city level and rescaled the daily flows from 29 January 2014 by the mean of the daily flow for 26–28 January period, aligning with the date of Wuhan's lockdown and the 2020 Chinese New Year holiday period. Then, the rescaled first dataset for 2020 under interventions was compared with the 2014 dataset to derive the percentage of travel decline for each city. The percentages for cities were averaged by day to preliminarily quantify the intensity of contact reduction in China under NPIs (Supplementary Table 2), as the policies of travel restriction and social distancing measures were implemented and occurred at the same time across the country.

We also collated data of the first case reported by county across mainland China to measure the delay from illness to case report as a reference of the improved timeliness of case identification, isolation and reporting during the outbreak (Supplementary Table 1). The daily number of COVID-19 cases by date of illness onset in the city of Wuhan, Hubei province and other provinces as of 13 February 2020 were used to further validate the epidemic curves estimated in this study across time. There was an abnormal increase of cases in Wuhan and Hubei province on 1 February 2020, on the basis of the date of illness onset[2]. We interpolated the number on 1 February 2020 by using the mean of numbers of cases reported on 31 January and 2 February 2020 in the epidemic curve. The number of cases reported by city across mainland China as of 29 February 2020 was used to define the predictability of our model across space. These case data were collated from the websites of national and local health authorities, news media and publications[2,3,31] (Supplementary Information).

Data analysis

We constructed a travel-network-based SEIR modelling framework (BEARmod) for before-and-after comparable analyses on the efficacy of NPIs. This model was extended from a typical SEIR model to specifically incorporate movement between locations that varied with each time step. In this model, each city was represented in the model as a separate subpopulation, with its own susceptible (S), exposed (E), infected (I) and recovered/removed (R) populations.

Exposure, infection and recovery

During each time step, infected people first recovered or were removed at an average rate r, where r was equal to the inverse of the average infectious period, and removal represents self-isolation and effective removal from the population as a potential transmitter of disease. This was incorporated as a Bernoulli trial for each infected person with a probability of recovering of $1 - \exp(-r)$. We used the median of time lags from illness onset to reported case as a proxy of the average infectious period, indicating the improving identification and isolation of cases under improved interventions (Supplementary Table 1). Then, the model converted exposed people to infectious by similarly incorporating a Bernoulli trial for each exposed individual, where the daily probability of becoming infectious $1 - \exp(-\varepsilon)$, where ε was the inverse of the average time spent exposed but not infectious, on the basis of the estimated incubation period (5.2 days, 95% confidence interval (CI) 4.1–7.0)[5]. Finally, to end the exposure, infection and recovery step of the model, the number of newly exposed people was calculated for each city on the basis of the number of infectious people in the city (I_i) and the average number of daily contacts that lead to transmission that each infectious person has (c). We simulated the number of exposed individuals in a patch on a given day through a random draw from a Poisson distribution for each infectious person, in which the mean number of new infections per person was c, which was then multiplied by the fraction of people in the city that were susceptible. We calculated the daily contact rate c using the basic reproduction rate that has been calculated in other studies ($R_0 = 2.2$ (95% CI 1.4–3.9)) divided by the average days (5.8, 95% CI 4.3–7.5) from onset to first medical visit and isolation[5], weighted by the relative level of daily contact where relevant, based on the Baidu movement data (Supplementary Table 2). Because simulation runs were not extended beyond five months, we did not include the addition of new susceptible people, or the conversion of recovered people back to susceptible.

The infection processes within each patch therefore approximate the following deterministic, continuous-time model, where c and r varied through time:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = S - c\frac{SI}{N}$$

$$\frac{\mathrm{d}E}{\mathrm{d}t} = c\frac{SI}{N} - \varepsilon E$$

$$\frac{dI}{dt} = \varepsilon E - rI$$

$$\frac{dR}{dt} = rI$$

Movement

After the model completed the infection-related processes, we moved infectious people between cities. To do this, we moved infected people from their current location to each possible destination (including remaining in the same place) using Bernoulli trials for each infected person, and each possible destination city. We parameterized the probability of moving from city i to city j (p_{ij}), which was equal to the proportion of smartphone users who went from city i to city j in the corresponding day from the Baidu dataset in 2015, accounting for the travel restrictions in 2020. This included modelling the numbers of people who stayed in the same location using p_{ii}, the proportion of users who did not move to a new location on that day. This allowed us to incorporate variance in the actual composition of travellers (infected versus non-infected), but because movement numbers were generated independently, it was possible for the number of infected people who stayed and the number who move in each patch to exceed or be fewer than the number of infected people in the patch. As we only wanted to incorporate variance into relative patterns of movement and not absolute numbers (particularly because the underlying values are proportions of people who moved and therefore cannot influence the total numbers of people infected), in any case in which the number of infected people who moved and the number who stayed differed from the total number of infected people in the origin patch, we rescaled values to the total number of infected people. Rescaling in this way meant the variance introduced by the Bernoulli trials could only influence relative movement patterns, and not actual numbers of infected people. Further, because we explicitly model the number of stayers in the same way as movers, rescaling should not introduce any bias in terms of the final relative movement patterns.

Through this model, stochasticity in the numbers and in the places with COVID-19 infections appears between simulation runs owing to variance in numbers of people becoming exposed, infectious and removed/recovered, as well as variance in numbers of people moving from one city to another. By modelling the COVID-19 epidemic in this way, we could simulate the incidence of COVID-19 cases, accounting for variance in recovery, infection and movement across many simulation runs (1,000). In addition, this allowed for us to account for uncertainty in contact rates after NPIs were implemented or lifted.

Simulation runs

Using this model, we quantified how the transmission of COVID-19 varied with different intervention scenarios and timings, as well as the potential of further transmission after the lifting of travel restrictions and contact distancing measures on 17 February 2020. As the earliest date of illness onset in cases was 2 December 2020 (ref. [3]), considering the underreporting of cases and the delay from infection to onset and identification of this novel virus, we started our simulations by infecting five people in Wuhan on 1 December 2019 and propagating the epidemic through time, varying factors including the timing and types of interventions used, assumed contact and recovery rates, and movement. We initially infected five people as a minimum number of infected people that prevented stochastic extinction of the epidemic during the initial days of simulation, and found no significant difference after three months, over simulation runs that started with three, five and eight people initially infected (though with three people initially infected, 50% of runs led to zero cases over the first week of simulation). When using data from other years we fixed the simulation dates around Chinese New Year and adjusted the start date of the epidemic accordingly.

The estimates of the model for the outbreak under NPIs as the baseline scenario were compared with reported COVID-19 cases across time and space. The sensitivity and specificity were also calculated to examine the performance of the model in predicting the occurrence of COVID-19 cases at the city level across China. The relative effects of NPIs were quantitatively assessed by comparing estimates of cases under various NPIs and timings with that of the baseline scenario. We also conducted a series of sensitivity analyses to understand the effect that changing epidemiological parameters had on the estimates and uncertainties of intervention efficacy. The software R v.3.6.1 (R Foundation for Statistical Computing) was used for data collation and analyses.

Ethical approval

Ethical clearance for collecting and using secondary data in this study was granted by the institutional review board of the University of Southampton (no. 48002). All data were supplied and analysed in an anonymous format, without access to personal identifying information.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The data on the number of cases of COVID-19 reported by county, city and province across China are available from the data sources listed in the Supplementary Information, and the average days from illness onset to reporting of the first case by each county used in the modelling are detailed in Supplementary Table 1. The mobile phone datasets analysed during the current study are not publicly available as this would compromise the agreement with the data provider; however, information on the process of requesting access to the data that support the findings of this study is available from S.L., and the data on travel and contact reductions that were derived from the datasets and used in our model are detailed in Supplementary Table 2.

Code availability

The code for the model built in this study has been made openly available for further use at https://github.com/wpgp/BEARmod.

28. Wang, X., Liu, C., Mao, W., Hu, Z. & Gu, L. Tracing the largest seasonal migration on Earth. Preprint at https://arxiv.org/abs/1411.0983 (2014).
29. National Bureau of Statistics of China. *China Statistical Yearbook 2014* http://www.stats.gov.cn/english/Statisticaldata/AnnualData/ (2020).
30. Kraemer, M. U. G. et al. Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. *Nat. Microbiol.* **4**, 854–863 (2019).
31. Zhang, J. et al. Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside Hubei province, China: a descriptive and modelling study. *Lancet Infect. Dis.* **20**, 793–802 (2020).

Acknowledgements We thank staff members at disease control institutions, hospitals and health administrations in areas across China in which outbreaks of COVID-19 occurred for field investigation, administration and data collection; Baidu Inc. for sharing population movement data; and Y. Zhu and S. Lai for collating online data. This study was supported by grants from the Bill and Melinda Gates Foundation (OPP1134076, OPP1195154); the European Union Horizon 2020 (MOOD 874850); the National Natural Science Fund of China (81773498); and the National Science and Technology Major Project of China (2016ZX10004222-009). A.J.T. is supported by funding from the Bill and Melinda Gates Foundation (OPP1106427, OPP1032350, OPP1134076, OPP1094793), the Clinton Health Access Initiative, the UK Department for International Development (DFID) and the Wellcome Trust (106866/Z/15/Z, 204613/Z/16/Z); H.Y. is supported by funding from the National Natural Science Fund for Distinguished Young Scholars of China (no. 81525023) and the Program of Shanghai Academic/Technology Research Leader (no. 18XD1400300); and O.P. is supported by funding from the National Science Foundation, USA (no. 1816075). The funders of the study had no role in study design, data collection, data analysis, data interpretation or writing of the report.

Author contributions S.L. designed the study, built the model, collected data, finalized the analysis, interpreted the findings and wrote the manuscript. N.W.R. built the model, analysed data, interpreted the findings and wrote the manuscript. L.Z. collected data, interpreted the

findings, commented on and revised drafts of the manuscript. J.R.F., O.P. and W.L. built the model, commented on and revised drafts of the manuscript. C.Z. collected data, interpreted the findings and commented on and revised drafts of the manuscript. A.J.T. interpreted the findings and revised drafts of the manuscript. A.W., M.S., X.D. and H.Y. interpreted the findings and commented on and revised drafts of the manuscript. All authors read and approved the final manuscript. The corresponding authors had full access to all of the data in the study and had final responsibility for the decision to submit for publication.

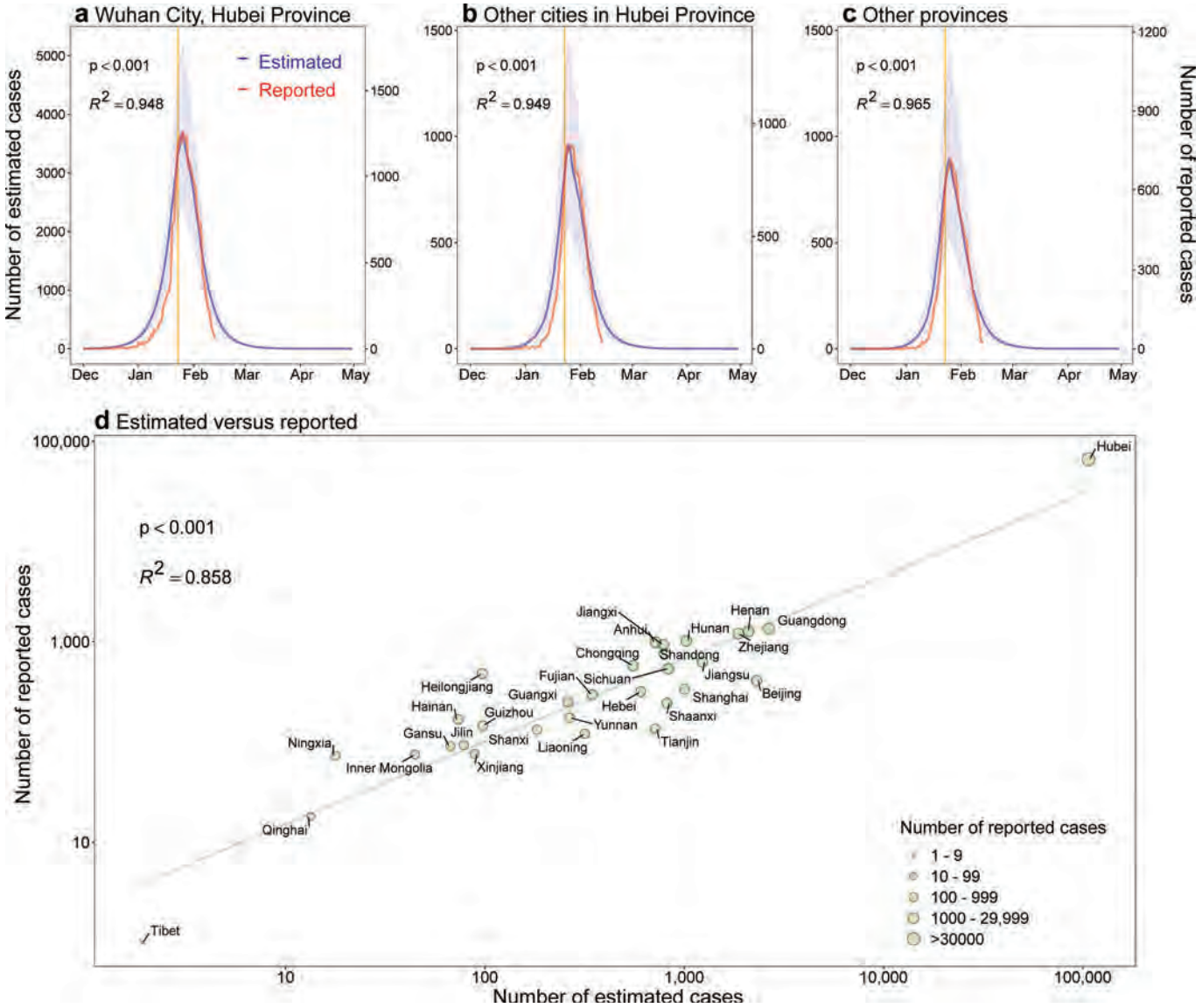
Competing interests H.Y. has received research funding from Sanofi Pasteur, GlaxoSmithKline, Yichang HEC Changjiang Pharmaceutical Company and Shanghai Roche Pharmaceutical Company. All other authors declare no competing interests.

Additional information

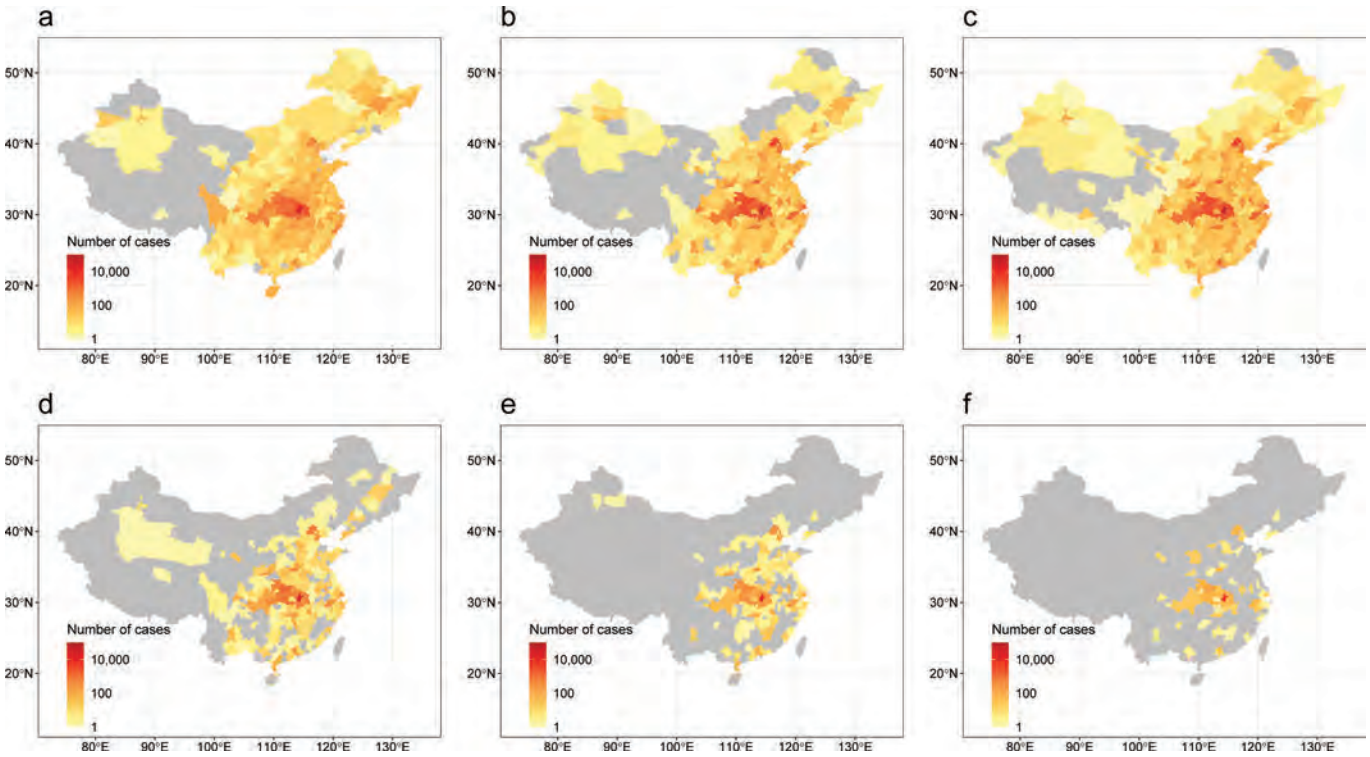
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2293-x.

Correspondence and requests for materials should be addressed to S.L., N.W.R. or A.J.T.

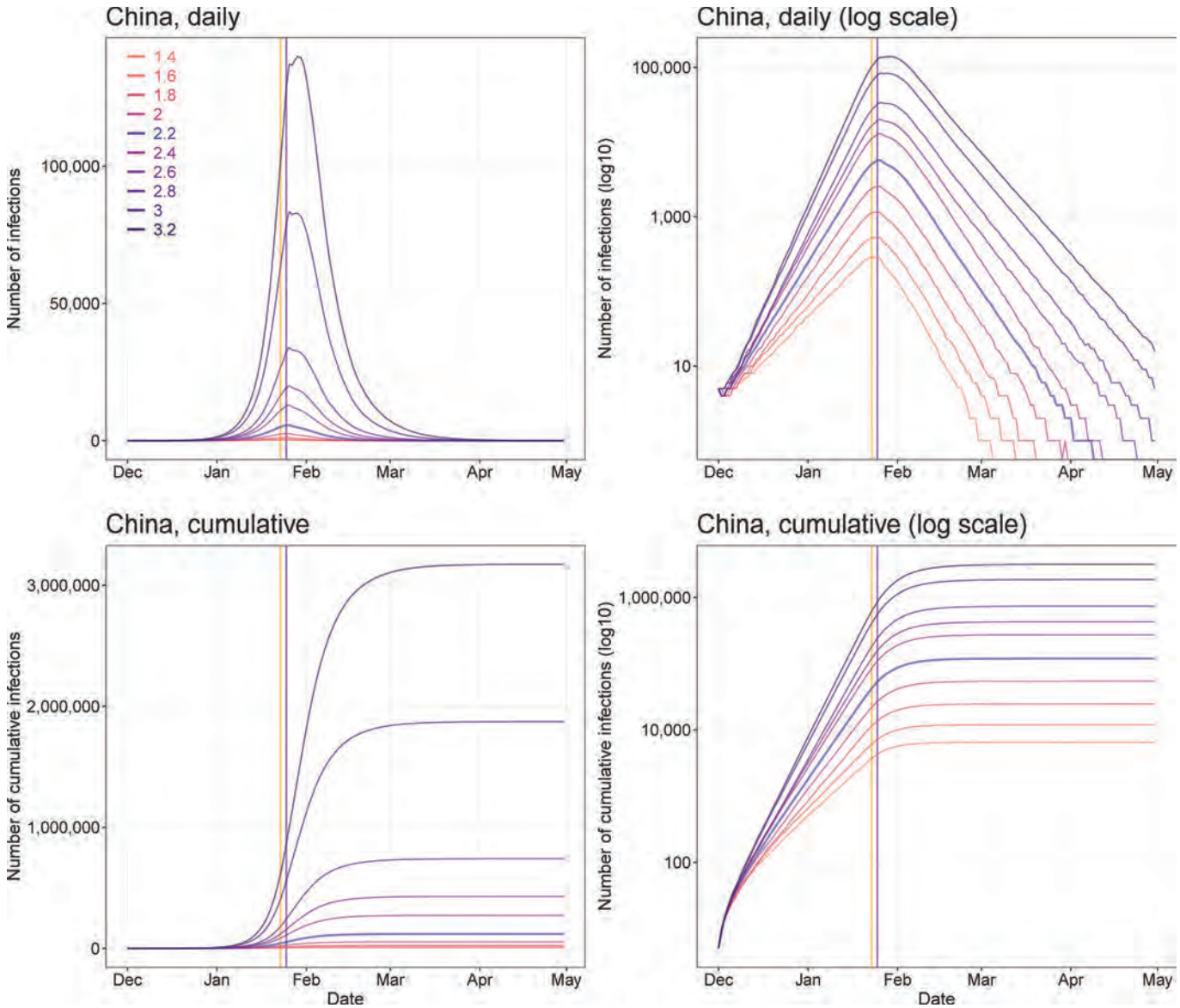
Peer review information *Nature* thanks Jukka-Pekka Onnela and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



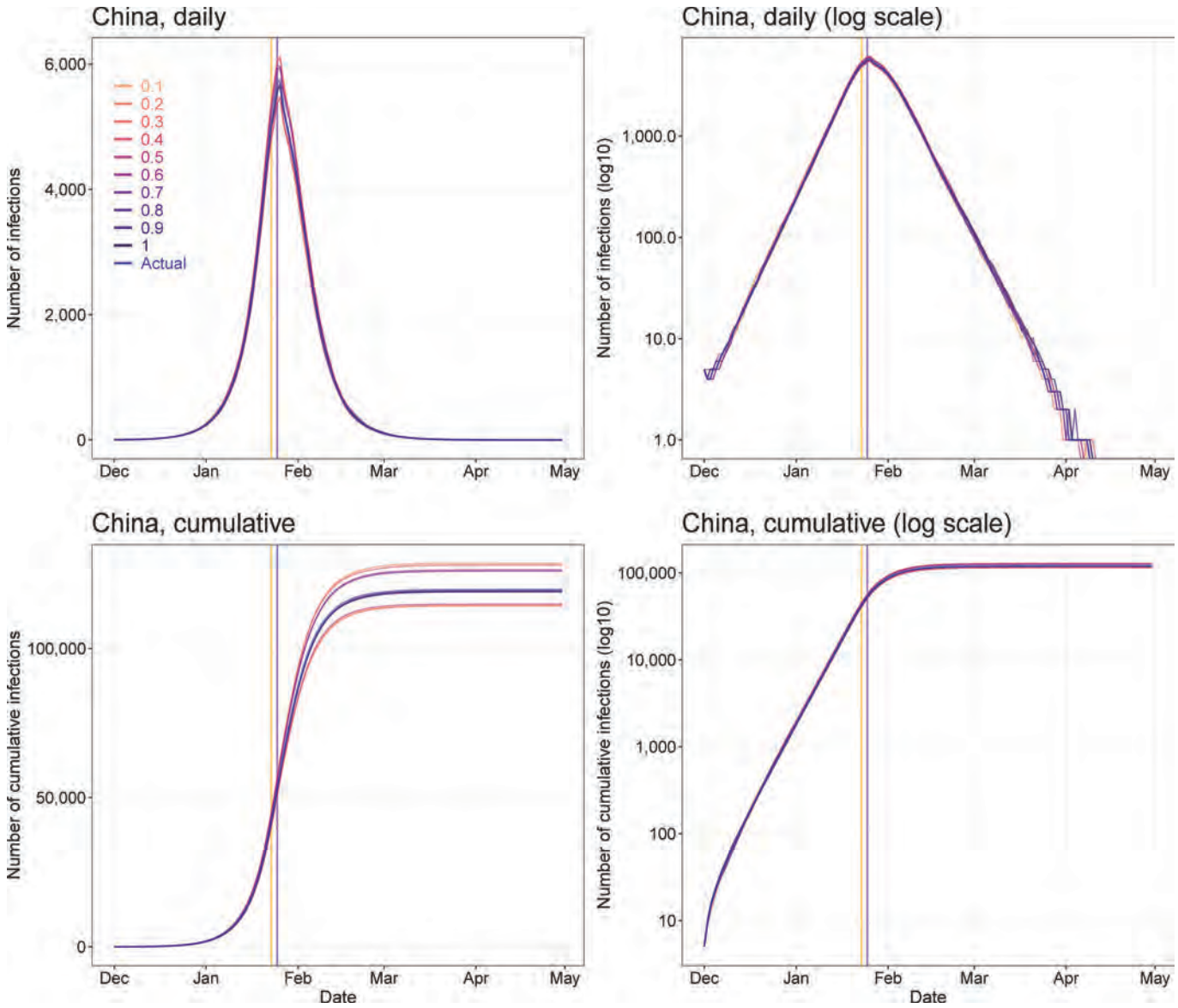
Extended Data Fig. 1 | Estimated and reported epidemic curves of the COVID-19 outbreak in mainland China. a, The city of Wuhan in Hubei province. **b**, Other cities in Hubei province. **c**, Thirty other provincial regions in mainland China. The orange vertical lines indicate the date on which Wuhan's lockdown began (23 January 2020). The estimated epidemic curves of COVID-19 cases show the median (dark-blue line) and IQR (light-blue shading) of estimates (1,000 simulations), and the Pearson's correlation between the median of daily estimates and the number of daily reported cases by region as of 13 February 2020 is also shown. **d**, The Pearson's correlation between the total number of estimated cases and the total number of reported cases by province as of 29 February 2020. *P* value calculated by two-sided *t*-test.



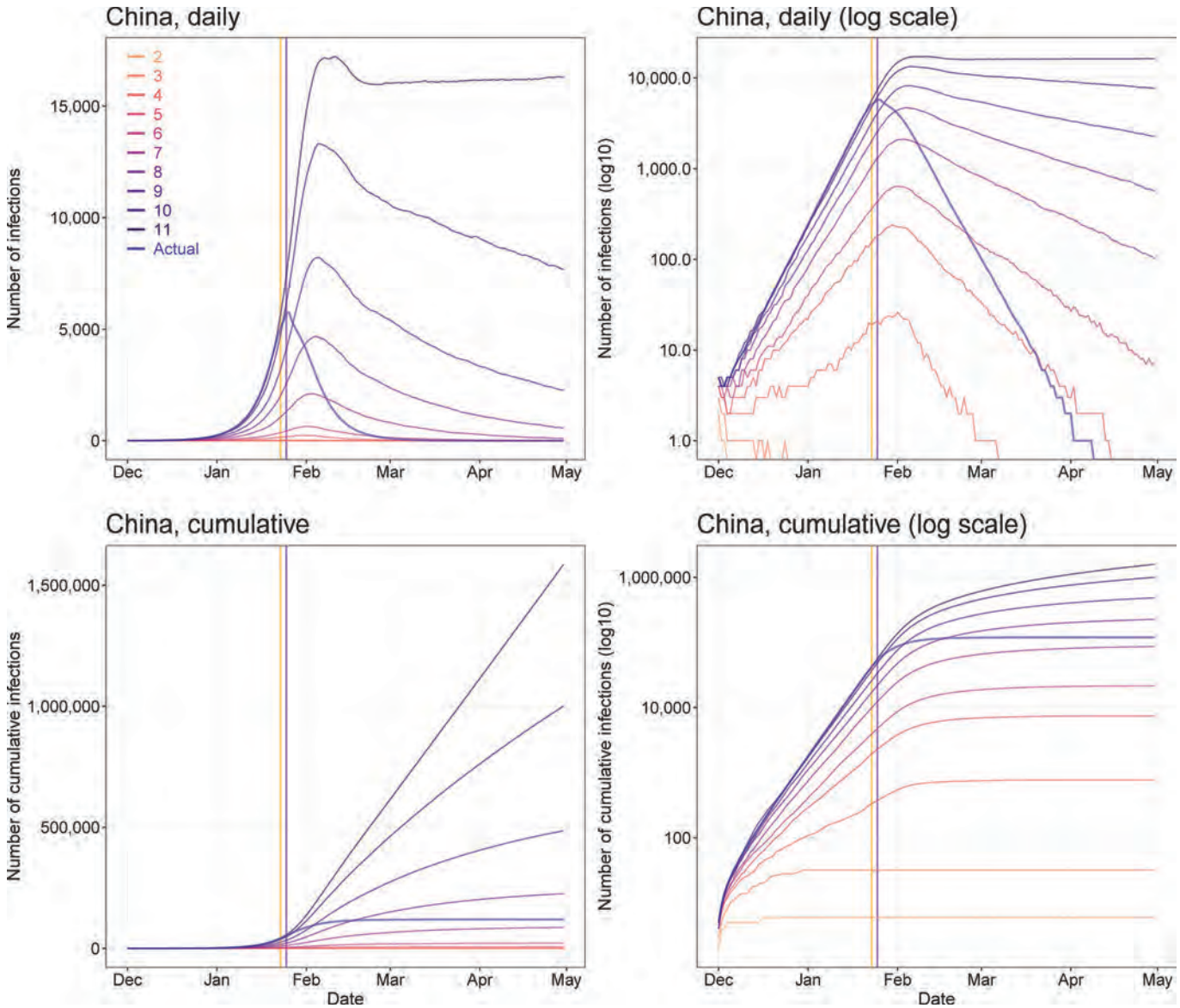
Extended Data Fig. 2 | Areas affected by COVID-19 in mainland China under various intervention timings. a, A total of 308 cities reported COVID-19 cases, on the basis of data obtained from national and local health authorities, as of 29 February 2020. **b**, Affected areas (298 cities) estimated by models under interventions implemented at actual timing. **c**, Estimated affected areas (326 cities) under interventions implemented at actual timing, but without intercity travel restrictions. **d**, Estimated affected areas (192 cities) under interventions implemented one week earlier than actual timing. **e**, Estimated affected areas (130 cities) under interventions implemented two weeks earlier than actual timing. **f**, Estimated affected areas (61 cities) under interventions implemented three weeks earlier than actual timing. The administrative boundary maps were obtained from the National Platform of Common Geospatial Information Services of China (www.tianditu.gov.cn).



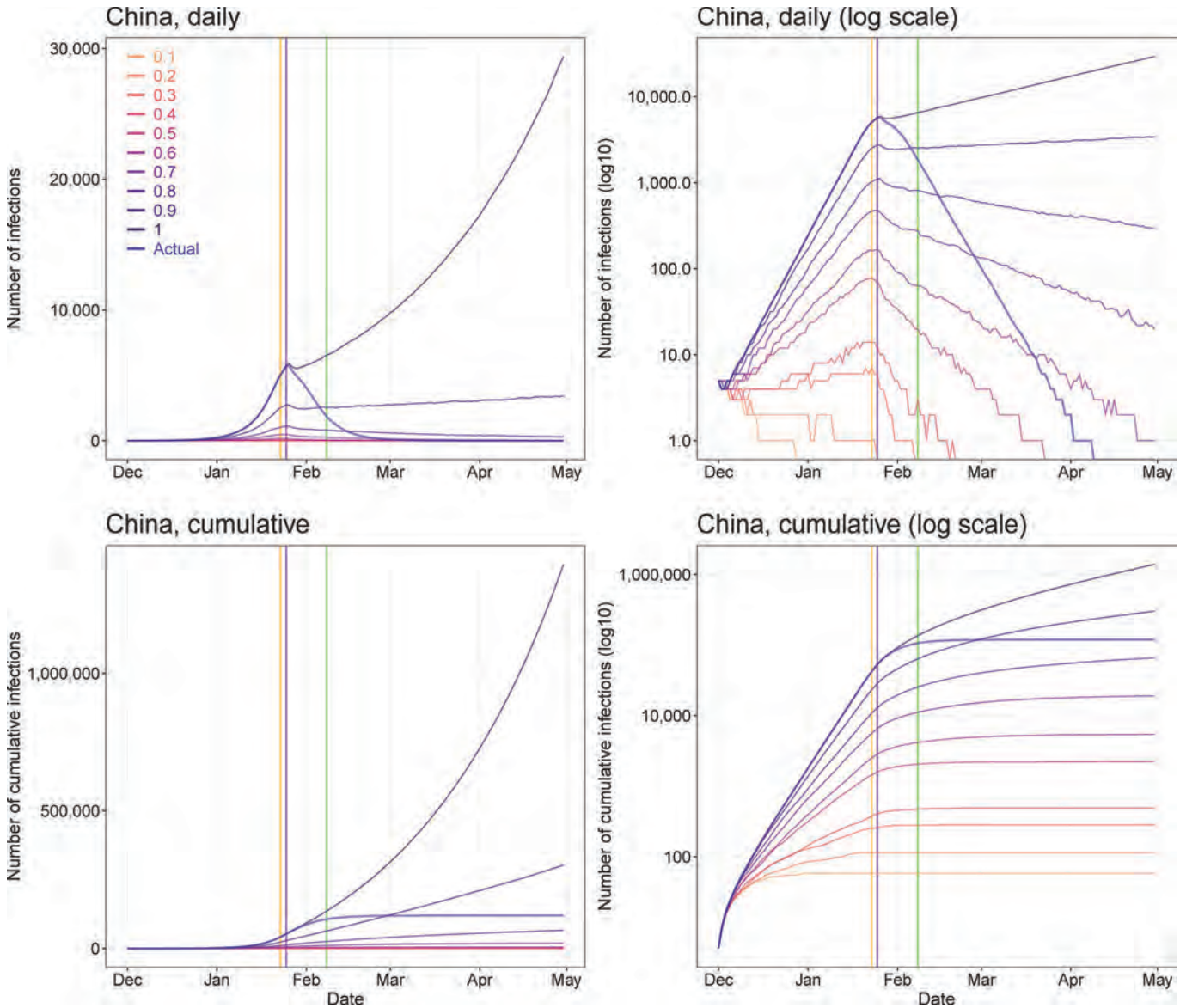
Extended Data Fig. 3 | Sensitivity of estimates of COVID-19 epidemics for various values of R_0. All other parameters, NPIs and input data were the same as the baseline model with $R_0 = 2.2$. Orange vertical lines indicate the date on which the lockdown of Wuhan began (23 January 2020); purple vertical lines indicate the date on which the Chinese New Year began (25 January 2020).



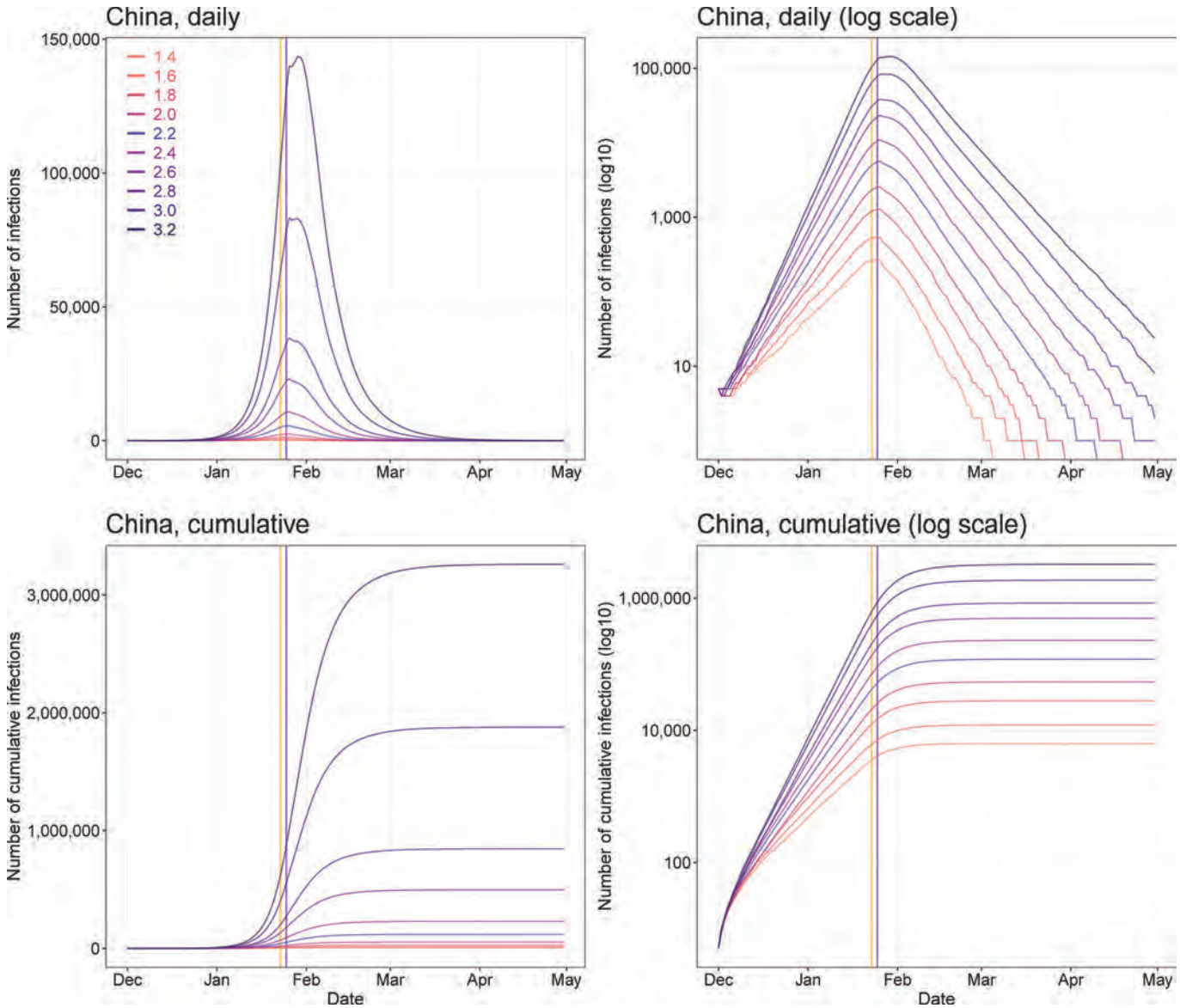
Extended Data Fig. 4 | Sensitivity of estimates of COVID-19 epidemics for various levels of intercity travel restrictions from 23 January 2020. All other parameters, NPIs and input data were the same as the baseline model with $R_0 = 2.2$. The actual percentages of intercity travel restrictions changed day by day across cities in China (0.1 indicates a 90% reduction from normal travel; 1 indicates no travel restrictions). Orange vertical lines indicate the date on which the lockdown of Wuhan began (23 January 2020); purple vertical lines indicate the date on which the Chinese New Year began (25 January 2020).



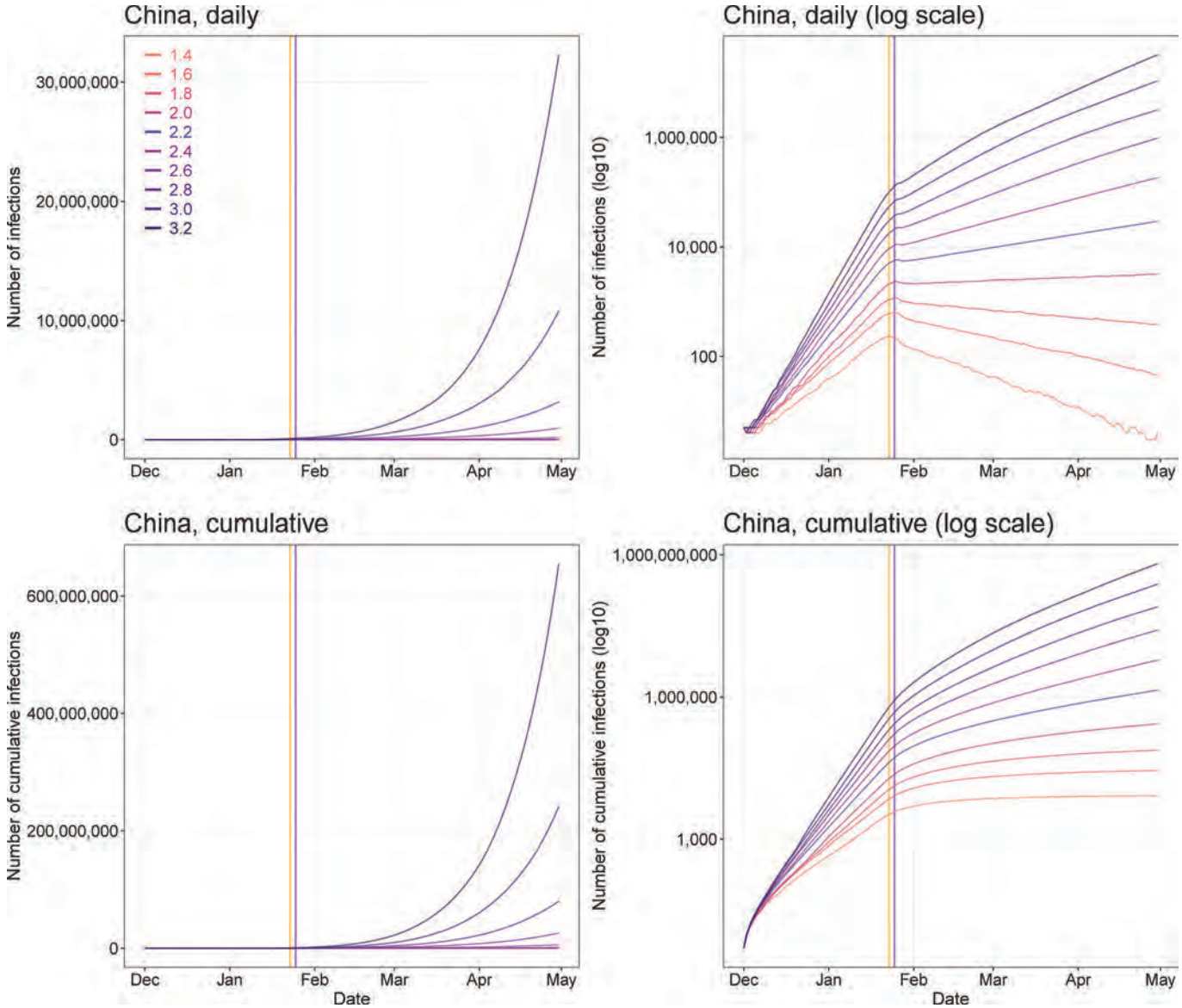
Extended Data Fig. 5 | Sensitivity of estimates of COVID-19 epidemics for various numbers of days from illness onset to report or isolation. All other parameters, NPIs and input data were the same as the baseline model with $R_0 = 2.2$. The actual delays of illness onset to report or isolation changed day by day (Supplementary Table 2). Orange vertical lines indicate the date on which the lockdown of Wuhan began (23 January 2020); purple vertical lines indicate the date on which the Chinese New Year began (25 January 2020).



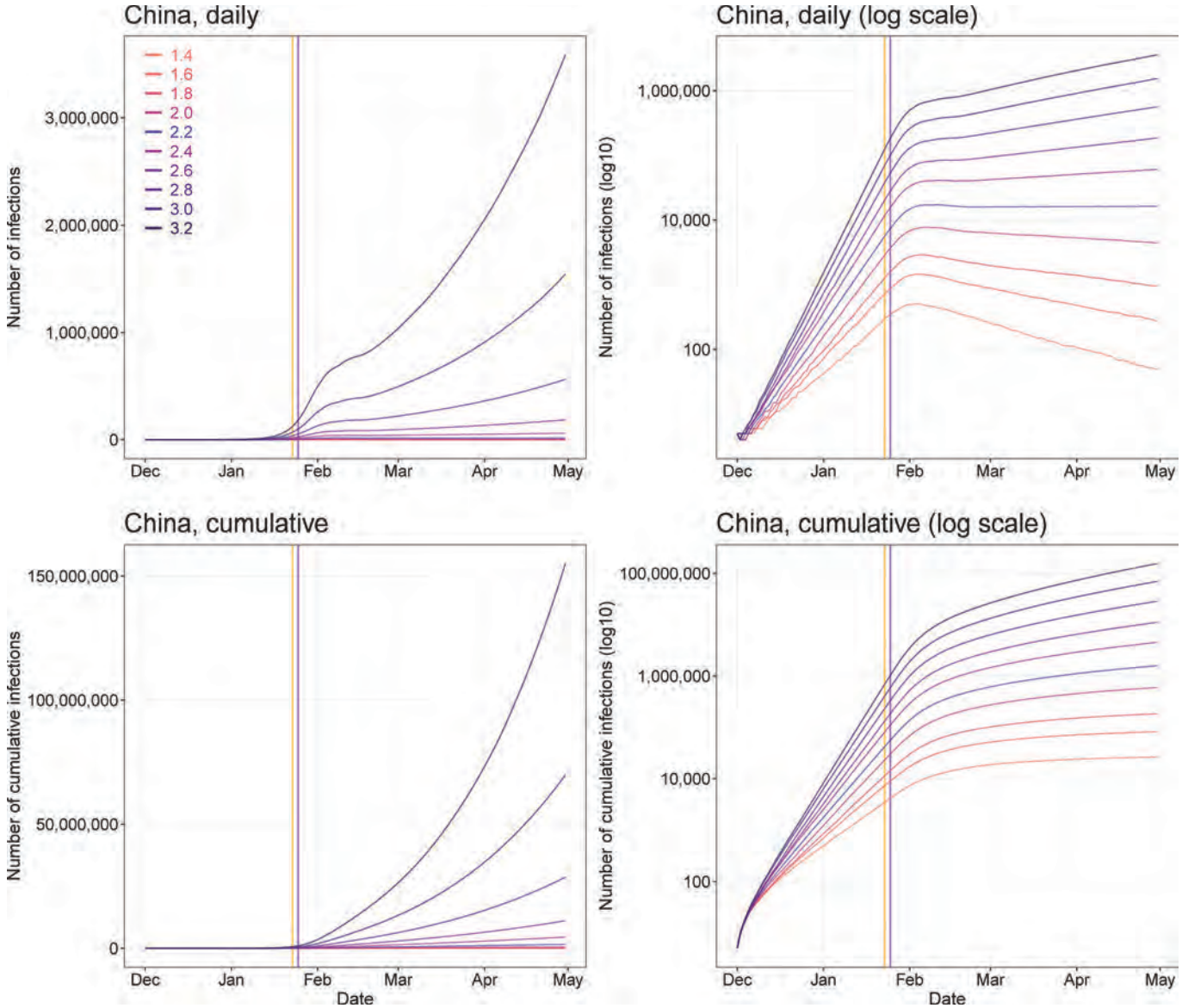
Extended Data Fig. 6 | Sensitivity of estimates of COVID-19 epidemics for various rates of contact. All other parameters, NPIs and input data were the same as the baseline model with $R_0 = 2.2$. The actual percentage of population contact (0.1 indicates 10% of usual contact, 1 means no contact restrictions) changed day by day across the country (Supplementary Table 1). Orange vertical lines indicate the date on which the lockdown of Wuhan began (23 January 2020); purple vertical lines indicate the date on which the Chinese New Year began (25 January 2020).



Extended Data Fig. 7 | Sensitivity of estimates of COVID-19 epidemics for various values of R_0 and without intercity travel restrictions. All other parameters, NPIs and input data were the same as the baseline model with $R_0 = 2.2$. Orange vertical lines indicate the date on which the lockdown of Wuhan began (23 January 2020); purple vertical lines indicate the date on which the Chinese New Year began (25 January 2020).



Extended Data Fig. 8 | Sensitivity of estimates of COVID-19 epidemics for various values of R_0 and without contact restrictions within cities. All other parameters, NPIs and input data were the same as the baseline model with $R_0 = 2.2$. Orange vertical lines indicate the date on which the lockdown of Wuhan began (23 January 2020); purple vertical lines indicate the date on which the Chinese New Year began (25 January 2020).



Extended Data Fig. 9 | Sensitivity of estimates of COVID-19 epidemics for various values of R_0 and without improved timeliness of case detection and isolation. The delay from illness onset to detection and isolation was set as a constant of 11 days, which was the average delay from 16–18 January 2020. All other parameters, NPIs and input data were the same as the baseline model with $R_0 = 2.2$. Orange vertical lines indicate the date on which the lockdown of Wuhan began (23 January 2020); purple vertical lines indicate the date on which the Chinese New Year began (25 January 2020).

Extended Data Table 1 | Reports and estimates of COVID-19 cases in mainland China as of 29 February 2020

Interventions and timing	Wuhan City, Hubei Province	Other cities in Hubei Province	Other provinces	Mainland China
Under current non-pharmaceutical interventions (NPIs)				
No. of cases reported (%)[a]	49,122 (62)	17,785 (22)	12,917 (16)	79,824 (100)
Estimates of cases (%)	78,910 (69)	18,503 (16)	16,912 (15)	114,325 (100)
Interquartile range	51,952-111,280	11,029-28,685	9,499-27,033	76,776-164,576
Dates of estimated peak	Jan 25-27	Jan 24-26	Jan 24-26	Jan 25-27
Interval between NPIs and epidemic peak[b]	7 days	6 days	6 days	7 days
Percentage (%) of cases that could have been prevented with earlier interventions				
One week ahead	61 (45-79)	71 (55-86)	78 (62-90)	66 (50-82)
Two weeks ahead	84 (78-89)	90 (82-94)	91 (84-95)	86 (81-90)
Three weeks ahead	94 (92-96)	97 (95-99)	98 (97-99)	95 (93-97)
Estimated relative no. of cases with later interventions [c]				
One week delay	2.4 (1.6-3.5)	3.1 (1.8-4.6)	3.3 (2-5.4)	2.6 (1.8-3.8)
Two weeks delay	5.8 (4.0-8.6)	8.6 (5.3-12.8)	9.4 (6.1-14.6)	6.7 (4.6-10.0)
Three weeks delay	15.1 (9-21.1)	22.6 (13.5-33.9)	27.9 (17.5-42.8)	17.6 (11.2-25.5)
Estimated relative no. of cases under various NPIs [c]				
Without inter-city travel restriction	1.0 (0.6-1.3)	1.1 (0.7-1.7)	1.1 (0.7-1.7)	1.0 (0.6-1.4)
Without inner-city contact reduction	2.5 (1.7-3.7)	2.6 (1.5-4.2)	2.4 (1.2-4.0)	2.6 (1.7-3.7)
Without case early detection and isolation	5.0 (3.3-6.9)	5.6 (3.2-8.4)	5.1 (2.5-8.4)	5.0 (3.3-7.1)
Without all interventions above	51.4 (33.2-71.2)	91.6 (57.6-132.5)	124.7 (77.4-180)	67.3 (43.7-93.7)

[a]The reported data on COVID-19 cases were obtained from the Chinese National Health Commission as of 29 February 2020.
[b]The timeliness of case identification and reporting improved from 19 January 2020 and the travel restrictions and social distancing were implemented from 23 January 2020. We compared the peak dates by region with 19 January 2020 to define the interval from NPIs to epidemic peak.
[c]Referring to the median of estimates under actual interventions and timing.
The median and IQR of estimates are shown.

nature research

Corresponding author(s): Shengjie Lai

Last updated by author(s): Apr 17, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	R version 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria) was used to perform data collation and analyses.
Data analysis	R version 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria) was used to perform data collation and analyses. The model built by this study has been made openly available for further use at https://github.com/wpgp/BEARmod.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data of COVID-19 cases reported by county, city, and province across China are available from data sources detailed in the Supplementary, and the average days from illness onset to report of the first case by each county used in the modelling are detailed in Supplementary Table 2. The mobile phone datasets analysed during the current study are not publicly available since this would compromise the agreement with the data provider, but the information on the process of requesting access to the data that support the findings of this study are available from Dr Shengjie Lai (Shengjie.Lai@soton.ac.uk), and the data of travel and contact reductions derived from the datasets and used in our model are detailed in Supplementary Table 1.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Quantitative observational and modelling study
Research sample	COVID-19 cases reported across mainland China as of February 29, 2020 were included in this study. As public awareness and enhanced case searching remained high throughout the study period, a high proportion of cases with symptoms was likely to have been detected, with nearly all reported cases eventually subjected to laboratory testing. However, the reported data of COVID-19 cases might not include asymptomatic and mild infections, and our model may have underestimated the total number of infections. The data on COVID-19 cases reported by county, city, and province across China are available from the data sources listed in the Supplementary Information File 3. This study also used population movement data across the country in 2020 and previous years, obtained from Baidu Location-based service. However, coverage biases of smartphone and Baidu users in population likely exist. Though a high percentage of the population owns smartphones in China, the mobile user group still does not cover specific subgroups of the population, particularly children. Therefore, our population movement data may provide an incomplete picture, and differences between the characteristics of smartphone owners and non-owners may also bias estimates in this study. Additionally, the magnitude and patterns of movements could change year by year.
Sampling strategy	This study included the numbers of all COVID-19 cases reported across mainland China as of February 29, 2020. Population movement data on human mobility of all Baidu users across the country were obtained from Baidu location-based service in 2014-2015 and 2020.
Data collection	We collated data of the first case reported by county across mainland China to measure the delay from illness to case report as a reference of the improved timeliness of case identification, isolation and reporting during the outbreak (Supplementary Information File 1). The daily number of COVID-19 cases by date of illness onset in Wuhan City, Hubei Province and other provinces as of February 13, 2020 were used to further validate the epicurves estimated in this study across time. The number of cases reported by city across mainland China as of February 29 were used to define the predictability of our model across space. These case data were collated from the websites of national and local health authorities, news media, and publications (Supplementary Information File 3). The epidemiological parameters estimated for the early stage of the outbreak in Wuhan from previous study (reference #5) were initially collected used to parameterise the epidemic before widely implementing interventions. Three population movement datasets, obtained from Baidu location-based services, were used in this study: 1) daily relative outbound and inbound flow of mobile phone users for each prefecture-level city (340 cities in mainland China) in 2020; 2) historical relative movement matrix with daily total number of users at city level from December 26, 2014 to May 26, 2015, aligning with the 2020 Chinese new year holiday period; 3) daily population movements at county level (2862 counties in China) from January 26 through April 30, 2014, aligning with the 2015 and 2020 Chinese new year holiday period.
Timing	COVID-19 cases: December 2, 2019 - February 29, 2020. Three Baidu population movement datesets: 1) January 26, 2014- April 30, 2014; 2) December 26, 2014 - May 26, 2015; 3) January 1, 2020 - April 13, 2020.
Data exclusions	Before conducting this study, we already noticed that there was an abnormal increase of cases in Wuhan City and Hubei Province on February 1, 2020, based on the date of illness onset. The case definition has been adjusted and a large number of clinically diagnosed cases before laboratory confirmation have been retrospectively reported into the information system since 12 February. However, the spike on February 1 might not represent the actual infection patterns. We have discussed this issue and underlying causes (e.g. changes of definitions, reporting delay, system error, incorrect reporting of the onset date) with epidemiologists in China, but exact reasons remain unclear. Therefore, before comparing reported data with estimates in our study, we interpolated the number on February 1 by using the mean of the numbers of cases reported on January 31 and February 2 in the epicurves of Wuhan and Hubei Province.
Non-participation	As this study collected and used secondary data from disease surveillance and Baidu location-based service, we did not access to the raw data and we don't know how many participants dropped out/declined participation. However, the number of COVID-19 cases might be unreported as asymptomatic and mild infections exist. The mobile user group does not cover specific subgroups of the population, particularly children, and not all mobile owners use the Baidu location-based service. Therefore, our population movement data may provide an incomplete picture of movement of all population in China, and the spatiotemporal and demographic variations in the behaviour of phone users could have biased population distribution and travel estimates.
Randomization	We did not randomly sample COVID-19 cases and movement data from population. In our SEIR modelling framework, we conducted 1000 simulations to account for the uncertainty of estimates.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Exhibit

28

NATURE | NEWS

Engineered bat virus stirs debate over risky research

Lab-made coronavirus related to SARS can infect human cells.

Declan Butler

12 November 2015

An experiment that created a hybrid version of a bat coronavirus — one related to the virus that causes SARS (severe acute respiratory syndrome) — has triggered renewed debate over whether engineering lab variants of viruses with possible pandemic potential is worth the risks.

In an article published in *Nature Medicine*[1] on 9 November, scientists investigated a virus called SHC014, which is found in horseshoe bats in China. The researchers created a chimaeric virus, made up of a surface protein of SHC014 and the backbone of a SARS virus that had been adapted to grow in mice and to mimic human disease. The chimaera infected human airway cells — proving that the surface protein of SHC014 has the necessary structure to bind to a key receptor on the cells and to infect them. It also caused disease in mice, but did not kill them.

Although almost all coronaviruses isolated from bats have not been able to bind to the key human receptor, SHC014 is not the first that can do so. In 2013, researchers reported this ability for the first time in a different coronavirus isolated from the same bat population[2].

The findings reinforce suspicions that bat coronaviruses capable of directly infecting humans (rather than first needing to evolve in an intermediate animal host) may be more common than previously thought, the researchers say.

But other virologists question whether the information gleaned from the experiment justifies the potential risk. Although the extent of any risk is difficult to assess, Simon Wain-Hobson, a virologist at the Pasteur Institute in Paris, points out that the researchers have created a novel virus that “grows remarkably well” in human cells. “If the virus escaped, nobody could predict the trajectory,” he says.

Creation of a chimaera

The argument is essentially a rerun of the debate over whether to allow lab research that increases the virulence, ease of spread or host range of dangerous pathogens — what is known as ‘gain-of-function’ research. In October 2014, the US government imposed a moratorium on federal funding of such research on the viruses that cause SARS, influenza and MERS (Middle East respiratory syndrome, a deadly disease caused by a virus that sporadically jumps from camels to people).

The latest study was already under way before the US moratorium began, and the US National Institutes of Health (NIH) allowed it to proceed while it was under review by the agency, says Ralph Baric, an infectious-disease researcher at the University of North Carolina at Chapel Hill, a co-author of the study. The NIH eventually concluded that the work was not so risky as to fall under the moratorium, he says.

But Wain-Hobson disapproves of the study because, he says, it provides little benefit, and reveals little about the risk that the wild SHC014 virus in bats poses to humans.

Other experiments in the study show that the virus in wild bats would need to evolve to pose any threat to humans — a change that may never happen, although it cannot be ruled out. Baric and his team reconstructed the wild virus from its genome sequence and found that it grew poorly in human cell cultures and caused no significant disease in mice.

“The only impact of this work is the creation, in a lab, of a new, non-natural risk,” agrees Richard Ebright, a molecular biologist and biodefence expert at Rutgers University in Piscataway, New Jersey. Both Ebright and Wain-Hobson are long-standing critics of gain-of-function research.

In their paper, the study authors also concede that funders may think twice about allowing such experiments in the future. "Scientific review panels may deem similar studies building chimeric viruses based on circulating strains too risky to pursue," they write, adding that discussion is needed as to "whether these types of chimeric virus studies warrant further investigation versus the inherent risks involved”.

Useful research

But Baric and others say the research did have benefits. The study findings “move this virus from a candidate emerging pathogen to a clear and present danger”, says Peter Daszak, who co-authored the 2013 paper. Daszak is president of the EcoHealth Alliance, an international network of scientists, headquartered in New York City, that samples viruses from animals and people in emerging-diseases hotspots across the globe.

Studies testing hybrid viruses in human cell culture and animal models are limited in what they can say about the threat posed by a wild virus, Daszak agrees. But he argues that they can help indicate which pathogens should be prioritized for further research attention.

Without the experiments, says Baric, the SHC014 virus would still be seen as not a threat. Previously, scientists had believed, on the basis of molecular modelling and other studies, that it should not be able to infect human cells. The latest work shows that the virus has already overcome critical barriers, such as being able to latch onto human receptors and efficiently infect human airway cells, he says. “I don't think you can ignore that.” He plans to do further studies with the virus in non-human primates, which may yield data more relevant to humans.

Nature | doi:10.1038/nature.2015.18787

References

1. Menachery, V. D. *et al*. *Nature Med.* http://dx.doi.org/10.1038/nm.3985 (2015).
2. Ge, X.-Y. *et al*. *Nature* **503**, 535–538 (2013).

Exhibit 29

Asian Economic Policy Review

doi: 10.1111/aepr.12359 *Asian Economic Policy Review* (2022) 17, 114–135

How COVID-19 Medical Supply Shortages Led to Extraordinary Trade and Industrial Policy

Chad P. BOWN†

Peterson Institute for International Economics & CEPR

Early in the COVID-19 pandemic, a global shortage of hospital gowns, gloves, surgical masks, and respirators caused policymakers globally to panic. China increased imports and decreased exports of this personal protective equipment, removing supplies from world markets. Shortages led to European Union and US export controls as well as other extraordinary policy actions, including a US effort to reserve supplies manufactured in China by a US-headquartered multinational. By April 2020, China's exports had mostly resumed, and over the rest of the year its export volumes surged. But China's export prices also skyrocketed and remained elevated through 2020, reflecting severe and continued shortages. This paper explores these and other government actions, such as US trade war tariffs and US industrial policy in the form of over $1 billion of subsidies to build out its domestic personal protective equipment supply chain, as well as potential lessons for future pandemic preparedness and international policy cooperation.

Key words: COVID-19, export restriction, industrial policy, personal protective equipment, supply chain, tariff

JEL codes: F13

1. Introduction

The early days of the COVID-19 pandemic brought fear and panic to the world for many reasons. A global shortage of basic personal protective equipment (PPE) was an important one. Nowhere to be found were hospital gowns and gloves, surgical masks and respirators, goggles, and face shields. Health care workers needed them in higher volume to take care of the unending surge of sick patients. But suddenly so did many others whose jobs put them in close proximity to coworkers, customers, or vulnerable populations.

For helpful discussions, I thank Tom Bollyky, Lucian Cernat, Hal Hill, Soumaya Keynes, Cassey Lee, Jong-Wha Lee, Liza Lin, Sam Lowe, Hiroshi Mukunoki, Marcus Noland, John Polowczyk, Shujiro Urata, Prashant Yadav, and the journal's editors. Hexuan Li and Eva Zhang provided data and research assistance. Melina Kolb, William Melancon, and Oliver Ward assisted with graphics. All remaining errors are my own.

†Correspondence: Chad P. Bown, Peterson Institute for International Economics, 1750 Massachusetts Avenue NW, Washington, DC 20036, USA. Email: cbown@piie.com

For policymakers in the USA and Europe, the PPE shortage of early 2020 was stupefying. Trade, especially with China, has been accused of being a major source of the problem. Policymakers have launched investigations into how things went so wrong, demanding change as a result. For example, shortly after assuming office in January 2021, President Joe Biden issued an Executive Order, stating, "this will never happen again in the United States, period. We shouldn't have to rely on a foreign country—especially one that doesn't share our interests or our values—in order to protect and provide for our people during a national emergency."[1]

This paper clarifies what is known about trade in PPE during the pandemic for China, the European Union (EU), and the USA. It also explores a series of extraordinary policies affecting PPE during the pandemic, including trade war tariffs, export controls, directives that multinational corporations prioritize American sales from their foreign subsidiaries, and new US industrial policy – including over $1 billion of subsidies to expand capacity along its domestic PPE supply chain. The present paper describes implications for post-pandemic policy and international cooperation, and explains where additional data collection and research efforts are needed.

2. Background on PPE Production and Trade

PPE includes a range of items.[2] The focus here is on surgical masks and respirators as well as "protective garments" – a broad category that includes hazmat suits, as well as some hospital gowns. The analysis also touches on hospital gloves, as well as goggles, face shields, and medical shoe coverings.[3]

On the demand side, consumption of PPE can be characterized by large positive externalities. The social benefit of wearing PPE during the pandemic was much larger than the (substantial) private benefit, given both the devastating health effects of the disease and its transmissibility via airborne particles. For example, one back-of-the-envelope estimate indicated that the social value of each cloth mask worn by the American public was $3000–$6000, whereas each N-95 respirator worn by a hospital worker could "easily be more than a million dollars" (Abaluck *et al.*, 2020). The divergence between private and social benefits is one motivation for policy intervention.

On the supply side, the USA and EU had pre-pandemic domestic manufacturing for some items, but product-level production data are not yet publicly available to clarify how much. However, the existence of some local production can be inferred from a variety of sources.

For the EU, intra-EU trade (e.g. France exporting PPE to Italy) is possible only with domestic production. Furthermore, in a March 2020 policy announcement (described below), the European Commission (2020a) stated that "production of personal protective equipment such as mouth protection masks in the Union is currently concentrated in a limited number of Member States, namely the Czech Republic, France, Germany, and Poland."

Another source of production information is company announcements; 3M and Honeywell, for example, reported expansions to their N-95 respirator manufacturing

product lines during the pandemic (3M, 2020a; Honeywell, 2020). A US International Trade Commission investigation in mid-2020 also described anecdotal evidence from industry interviews (United States International Trade Commission [USITC], 2020).

For some products, however, there was apparently little pre-pandemic domestic production, at least in the USA. John Polowczyk, who led the US government's PPE Supply Chain Task Force from March 15 through November 2020, said "we made about 500 million nitrile gloves in America, pre-pandemic. [During the pandemic] we were using 1.8 billion a week. 500 million a year for manufacturing is not like you just get to put on another shift and make more gloves."[4]

Changes in the US domestic regulatory environment were also likely to have impacted PPE availability. For example, one agency (the National Institute for Occupational Safety and Health [NIOSH]) regulated the N-95 respirator for industrial use and another (the Food and Drug Administration [FDA]) regulated it for medical use. Before the pandemic, more than 95% of American N-95 respirator use was in industrial rather than medical settings, to protect workers from dust, chemicals, or other hazardous airborne particles (USITC, 2020; p. 89). (This use likely declined periodically throughout 2020, when lockdowns emerged). In March 2020, the FDA facilitated product availability by authorizing emergency use of NIOSH-approved N-95 respirators in medical settings.

However tempting, it is impossible at this stage to definitively attribute changes in trade flows during the pandemic to policy changes. That is because multiple determinants of domestic supply and demand – and thus imports and exports – were changing alongside many of the policy changes described next. As an example, for a net exporting country of PPE, increased demand for PPE because of a domestic coronavirus outbreak and decreased supply due to an industrial lockdown would each have the same impact – reducing export volumes – as a newly imposed export-restricting policy. Alternatively, relaxing the stringency of the regulatory environment might increase both domestic and foreign supply of N-95 respirators, but without knowing which was bigger, such a change would have an uncertain net effect on imports. In addition to trade data and an economic model, a rigorous assessment requires extremely detailed data on the domestic production and consumption of PPE before and during the pandemic, and these data are not yet publicly available at the level of disaggregation needed.

The following sections present stylized facts on PPE trade flows in light of several major policy actions, although even that effort is confounded by measurement challenges. For example, the most precisely defined pre-pandemic PPE product classifications often also included unrelated items in the tariff schedule (in examining changes over time, the assumption is that there was little pandemic-related change in demand for or supply of those other items). Furthermore, volumes are often measured in weight (e.g. kilograms), not more familiar units often referenced by policymakers, such as number of masks.

Before the pandemic, China was the top exporter of most of the products considered in this analysis (Figure 1). The exception was hospital gloves (Malaysia). The USA and EU applied relatively low most favored nation (MFN) import tariffs on these products.[5]

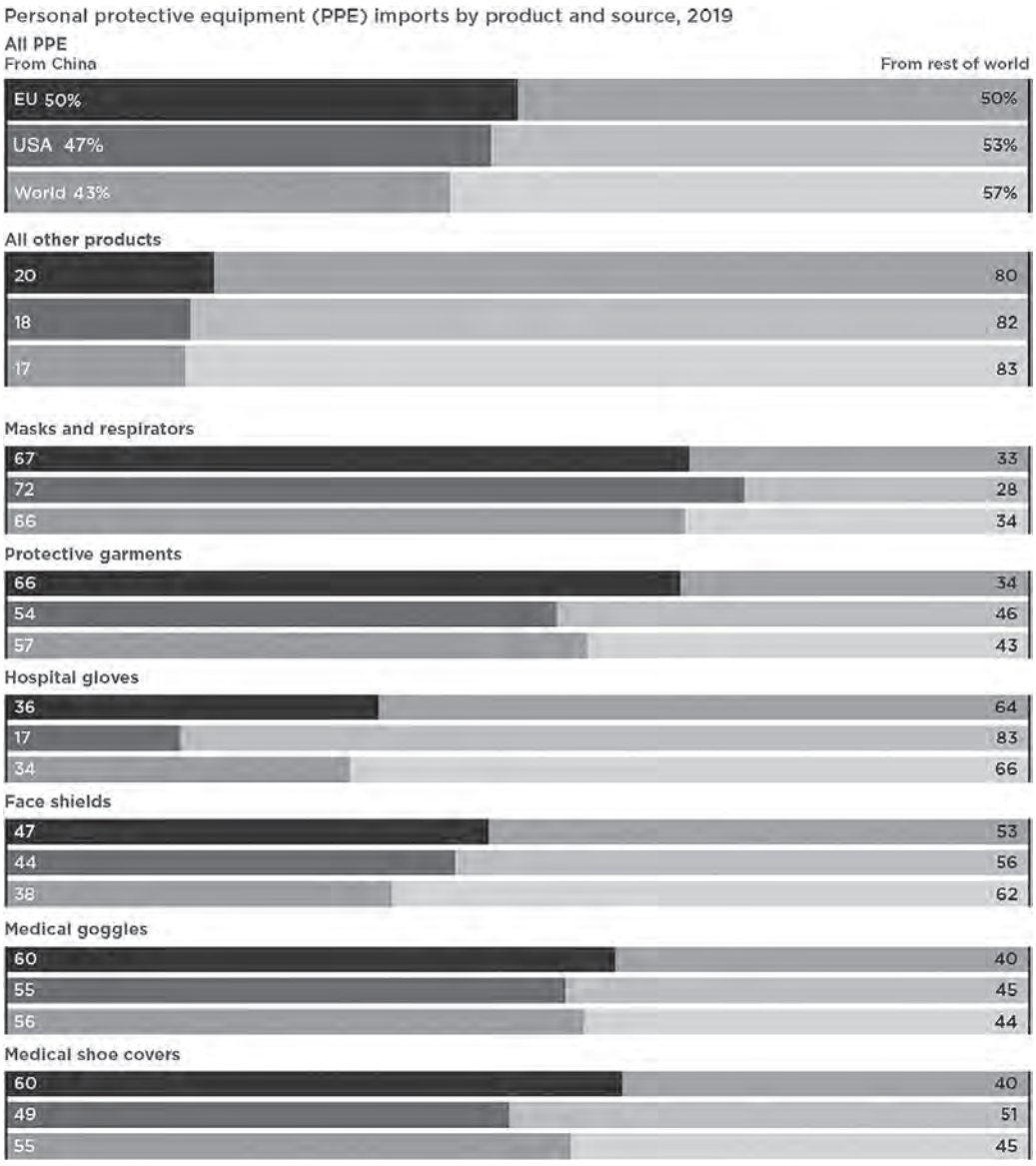


Figure 1 European Union (EU), USA, and world imports of personal protective equipment from China, 2019.

Source: Constructed by the author with data from US Census, Eurostat, and UN Comtrade.

Note: For the European Union, extra-EU imports only.

Supporting Information Table S1 provides a complete timeline of events discussed in the next three sections.

3. What Happened in China

In late December 2019, a novel coronavirus was discovered in the city of Wuhan in China's Hubei province. A month later, on January 30, 2020, the World Health

Organization (WHO) declared the COVID-19 outbreak a Public Health Emergency of International Concern. The Chinese government locked down parts of the economy, imposed travel restrictions, and even built entirely new hospitals from scratch. China found itself in desperate need of PPE.

In ernational markets at the time worked as expected from stress: China imported more, and exported less (Figure 2a).[6] The change in net exports for each product was dominated by China's reduction in exports. The magnitude of the decline in net exports in February 2020 dwarfed the similar seasonal reduction in 2018 and 2019, associated with the Chinese Lunar New Year (Figure 2b). Cumulating trade volumes over the first 3 months of 2020, China's exports of PPE were significantly lower than in the first quarter of 2019.[7] For example, export volumes were 12.5% lower for masks and respirators and 22.1% lower for protective garments.

Much of the decline in China's exports of protective garments, for example, can be traced directly to Hubei, the source of the outbreak – and of more than one-third of the country's total exports of protective garments in 2019. Hubei's export decline accounted for roughly 75% of the drop in China's total exports of protective garments in the first quarter of 2020.[8] For other products, the link between the COVID-19 shock and export concentration was less tight. For example, Zhejiang, Guangdong, Shanghai, and Jiangsu were the combined source of three quarters of China's exports of masks and respirators, products that also saw a significant export decline in the first quarter of 2020. While media reported China had also restricted PPE exports, the Chinese government denied the allegations (e.g. Hui, 2020).

By early March, the Chinese government announced a significant expansion of domestic PPE production. On March 6, the State Council (2020) indicated that China's daily output of protective clothing had increased from "less than 20,000 pieces in the early stage of the epidemic to the current 500,000 pieces. N-95 masks reached 1.6 million from 200,000, and ordinary masks reached 100 million."

China's net exports regained pre-pandemic (monthly) levels for most products by April 2020. Indeed, that month's mask and respirator exports were nearly double pre-pandemic levels, and exports of protective garments were 60% higher. Export volumes for most products remained elevated through the remainder of 2020.

But even this significant scaling up of export volumes was insufficient to satisfy exploding global demand (Figure 2b), and Chinese export prices (unit values) for most products rose even more than the volume increase. For masks and respirators, they were over 700% higher in April 2020 than before the pandemic, even with the doubling of volumes, and for protective garments they were more than 500% higher. For most products, export prices remained elevated through the end of 2020.

PPE scarcity and exploding prices generated a separate problem: counterfeit products. On April 10 the Chinese government responded by establishing a new system of quality controls for exports of various medical supplies, including nine PPE products.[9] One governmental concern was that a few bad actors could create large, negative reputational spillovers impacting the important Chinese PPE exporting industry.[10]

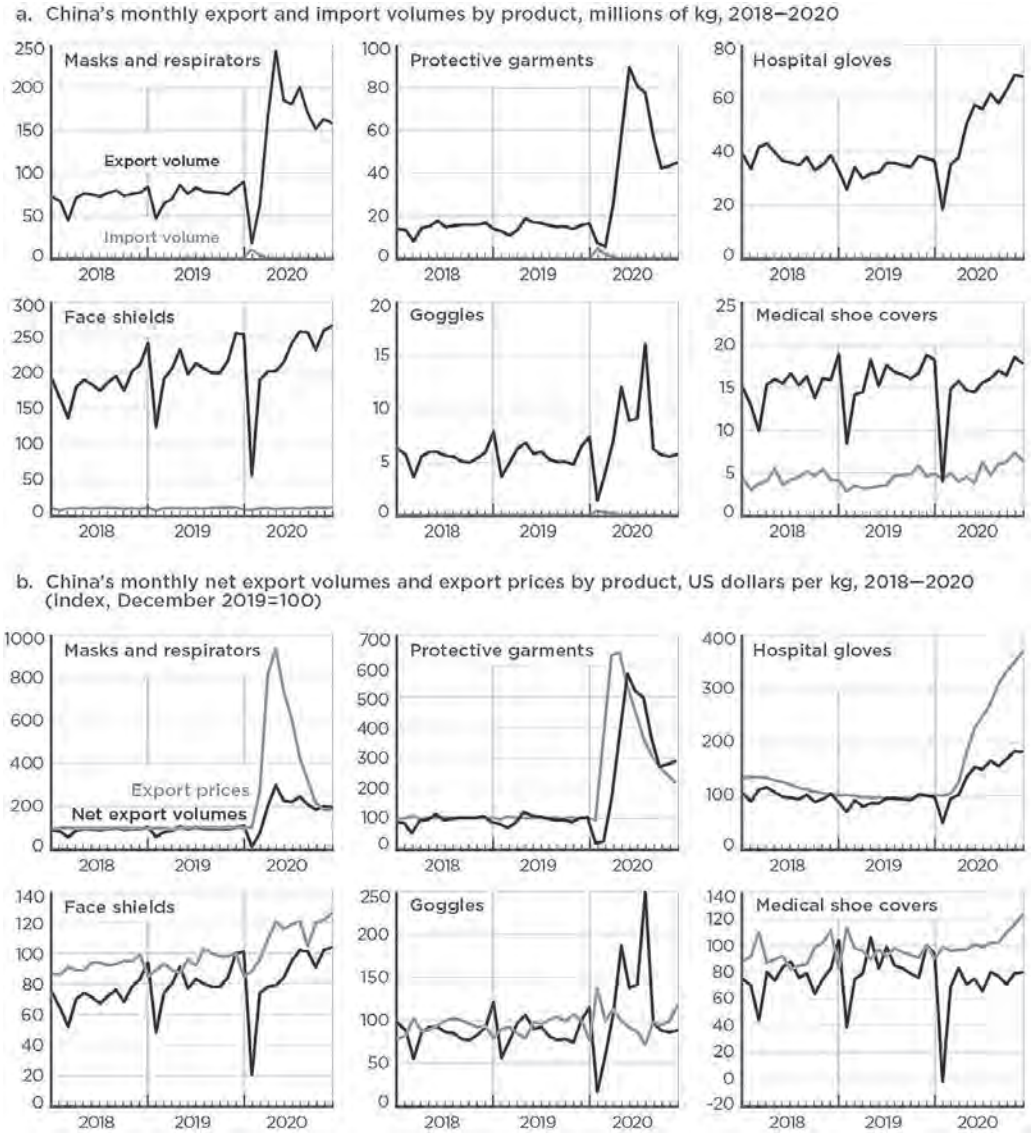


Figure 2 COVID-related drop and then surge in China's personal protective equipment net exports and prices.

Source: Chinese Customs

Note: Export price in trade unit values. For products with multiple eight-digit HS codes, the largest (by value in 2020) is shown. See Supporting Information Table S2

The US government quickly worried that Beijing was taking advantage of its market power and restricting exports for other, potentially political, reasons (see O'Keeffe *et al.*, 2020). Yet the trade data do not suggest that either China's export quality controls or US-China tensions affected China's PPE exports to the USA: the growth in these exports generally followed the same qualitative pattern as China's exports to the EU and the rest of the world.

Despite the considerable drop-off in the first quarter, China's PPE export performance over the rest of the year was stunning: the value of these PPE exports nearly quadrupled from $22.9 billion in 2019 to $88.1 billion in 2020 (Table 1). Relative to a year earlier, China's export volumes in the second through fourth quarters of 2020 were 130% higher for masks and respirators and 272% higher for protective garments. As terrible as things were early in 2020 when China's decline in net exports left many countries exposed, China's increasing exports over the rest of 2020 likely helped medical workers around the world save many lives.

4. What Happened in Europe

As the pandemic spread globally in early 2020, conditions in Europe began to deteriorate. In February, Italy experienced a spike in cases; Spain and other European countries also suffered, and policymakers panicked. On March 3, France requisitioned PPE for its health workers, and on March 4 Germany banned PPE exports. The French and German curbs applied even to exports destined for other EU member states, including Italy, which went into lockdown on March 9.

The export bans were also problematic because the countries imposing them were home to some of the EU's largest PPE production facilities. On March 15, the European Commission stepped in with a similarly unprecedented imposition of EU-wide export controls on PPE, in an attempt to get EU member states to free up shipments with each other (see, e.g., Bown, 2020b; European Commission, 2020a,b; Keynes, 2020).

Despite Europe's increasing needs, China's PPE exports to the EU declined in the first quarter of 2020, ending up 4–25% lower, depending on the product, relative to the same period in 2019 (Table 1). Even when China's export volumes recovered to pre-pandemic levels starting in April, prices skyrocketed, revealing the severity of the shortage. Compared to December 2019, China's prices of exports to the EU in April 2020 were 1250% higher for masks and respirators and 700% higher for protective garments.

For some products, Europe was not able to substitute imports from alternative suppliers.[11] For protective garments, for example, the decline in import volumes from China in the first quarter was accompanied by only a slight increase in intra-EU shipments, and imports from the rest of the world were flat. Imports of hospital gloves declined from all sources. In April, imports of most products began to accelerate, with the largest increases in imports from China, which continued over the last three quarters of 2020. EU import prices on most PPE also rose sharply, first from China and then from other sources. However, for most products, the price increase of imports from China was much higher than for imports from the rest of the world.

The products that the European Commission subjected to export controls on March 15 tell a mixed story. For many products, extra-EU export sales fell in March and April 2020, the period during which most of the export controls were in effect (Figure 3). However, it is difficult to disentangle how much of the export reduction resulted from EU policy, since other factors were changing at the same time. Internal EU demand for PPE was increasing, imports from China had fallen, and intra-EU exports for some products

Table 1 China's exports of PPE in 2019 and 2020, by product and destination

Trade values (billions, USD)

	In 2019			In 2020		
Product	Total	…to EU	…to USA	Total	…to EU	…to USA
Masks and respirators	5.4	1.0	2.2	53.8	17.3	14.8
Protective garments	0.9	0.3	0.4	10.8	2.6	2.7
Hospital gloves	1.0	0.1	0.5	3.9	0.6	2.0
Face shields	13.3	2.1	3.5	16.8	2.4	4.4
Goggles	1.4	0.3	0.4	1.9	0.4	0.5
Medical shoe covers	0.9	0.1	0.2	1.0	0.1	0.2
Total	22.9	3.9	7.1	88.1	23.4	24.6

Trade volumes, year-over-year percent changes (volume)†

	In January–March 2020			In April–December 2020			In 2020		
Product	Total	… to EU	…to USA	Total	…to EU	…to USA	Total	…to EU	…to USA
Masks and respirators	−12.5	−11.4	−18.3	130.0	183.1	77.9	99.1	140.4	58.0
Protective garments	−22.1	−22.9	−30.6	271.7	179.4	184.4	210.0	138.3	137.6
Hospital gloves	−3.0	−15.5	−5.4	68.0	171.8	59.9	51.7	120.7	46.0
Face shields	−10.2	−19.3	−10.0	10.7	3.5	19.4	6.0	−2.4	12.6
Goggles	−22.0	−24.5	−15.0	48.7	33.5	58.4	31.6	18.9	38.1
Medical shoe covers	−10.9	−4.4	−10.7	−2.7	3.3	3.7	−4.5	1.7	0.7

Trade prices (unit values), percent changes in Chinese export price†

	In April 2020 versus December 2019			At peak versus December 2019			In December 2020 versus December 2019		
Product	Total	… to EU	…to USA	Total	…to EU	…to USA	Total	…to EU	…to USA
Masks and respirators	720.8	1251.7	542.1	838.0	1251.7	778.1	85.6	149.3	51.3
	536.2	698.8	256.7	543.6	698.8	394.2	116.7	124.4	110.5

Trade prices (unit values), percent changes in Chinese export price†

Table 1 *continued*

Trade prices (unit values), percent changes in Chinese export price[†]

	In April 2020 versus December 2019			At peak versus December 2019			In December 2020 versus December 2019		
Product	Total	… to EU	…to USA	Total	…to EU	…to USA	Total	…to EU	…to USA
Protective garments									
Hospital gloves	24.8	66.9	15.5	270.4	262.0	279.5	270.4	262.0	279.5
Face shields	9.2	18.5	5.8	26.4	32.4	22.3	26.4	27.2	19.3
Goggles	12.6	57.3	−15.1	36.7	57.3	13.0	16.8	−0.4	5.5
Medical shoe covers	−3.4	−6.4	−0.3	23.3	41.5	35.4	23.3	37.9	28.3

Notes: Values defined using all HS08 codes for that product. Totals may not sum due to rounding.

[†]Volumes and prices (unit values) rely on only the top HS08 code by value in 2020. Percent changes in volume data are year over year for the relevant period. Price changes are month over month as indicated in the table.

EU, European Union; PPE, personal protective equipment.

(e.g. masks and respirators, protective garments) were increasing. Furthermore, domestic production may have been affected – at some points by lockdowns, at others by capacity expansion. The fact that exports did not surge after the EU export control regime expired suggests that low export volumes may not have been the result of policy but were dominated by these other factors, although it is impossible to say without more detailed production and consumption data.

Though the price of some EU PPE exports increased considerably (Figure 3b), the price increase for masks and respirators as well as protective garments was not nearly as high as for Chinese exports (see again Figure 2 and Table 1). This raises the question of whether the EU export monitoring system allocated PPE – in short supply globally – through a mechanism that was less responsive to price.

5. What Happened in the USA

The pandemic similarly hit the USA hard, beginning most famously in New York City, which declared a state of emergency on March 12. By early April, the Strategic National Stockpile for PPE, administered by the US Department of Health and Human Services (HHS), was essentially depleted. China's exports to the USA largely mimicked the European experience, declining in the first quarter of 2020 by 5–31% year-over-year, depending on the product (Table 1). The decrease in US imports from China was not accompanied by a comparable increase in imports from elsewhere.[12]

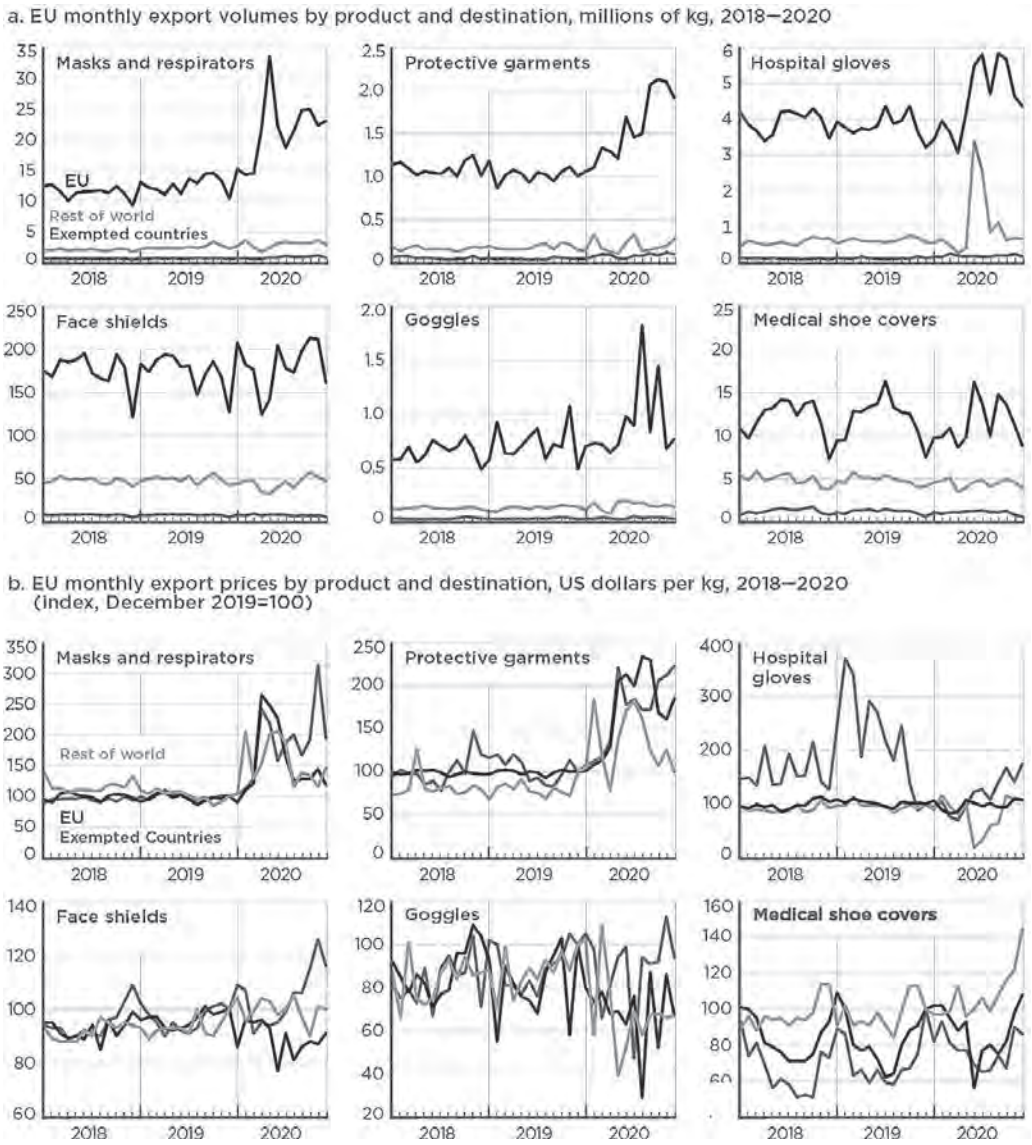


Figure 3 European Union (EU) personal protective equipment exports subject to export controls in March and April 2020.

Source: Eurostat.

Note: Export price in trade unit values. For products with multiple eight-digit HS codes, the largest (by value in 2020) is shown. See Supporting Information Table S2

Chinese exports to the USA also regained pre-pandemic levels by April and then increased considerably. Chinese export prices also skyrocketed, remaining high for much of the rest of 2020, reflecting continued shortages. China was the source of most of the increase in US import volumes in the second through fourth quarters of 2020 for most products. Imports of masks and respirators, as well as hospital gloves, also

began arriving in significantly increased quantities from Vietnam and Malaysia, respectively.

The volume of China's PPE exports to the USA in 2020 was somewhat remarkable, given that the US government sent mixed messages about whether it wanted imports of Chinese medical supplies. The Federal Emergency Management Agency (FEMA) created Project Airbridge to ship planeloads of PPE directly into the USA, beginning in late March, including from China. But statements from White House official Peter Navarro and Secretary of State Mike Pompeo, as well as President Trump's continued public references to the "Chinese virus," threatened to imperil the bilateral relationship during much of 2020. In addition, there were the US trade war tariffs.

5.1 Section 301 tariffs and US pandemic preparedness

The US administration began a trade war with China in 2018 that ultimately resulted in new US tariffs covering $335 billion, or two-thirds, of its goods imports from China. This included new tariffs on billions of dollars of imported medical equipment, despite warnings from experts that the duties could affect American preparedness for a future pandemic. When COVID-19 arrived, AdvaMed, an industry association, sent the office of the US Trade Representative (USTR) a letter on January 31 urging removal of the trade war tariffs on desperately needed medical supplies, including PPE. The administration stubbornly took many weeks to decide; for example, USTR did not grant temporary exclusions for masks and respirators until March 17.[13]

The trade war tariffs, implemented as part of the US administration's explicit policy goal of limiting imports, likely had a negative impact on US pandemic preparedness.[14] In the 4 months immediately following the September 2019 imposition of new tariffs, the year-over-year change in US imports from China was negative for four out of five PPE products facing those tariffs (Figure 4). By January 2020, for three of the most important product lines – face shields, masks and respirators, and protective garments – that absolute decline in imports was not offset by a commensurate increase in imports from elsewhere.

Overall, this suggests that the American health care system bought less from China and did not restock inventories from alternative foreign suppliers. With higher prices resulting from the tariffs, some American buyers may have also severed commercial relationships with Chinese suppliers that may have been difficult to restart in the midst of a pandemic.

5.2 The Defense Production Act arrangement with 3M's plants in China

In early April 2020, American PPE shortages had become so dire that the US administration invoked the Defense Production Act (DPA). One extraordinary element was its instruction to 3M to import 166.5 million respirators over April, May, and June from its plants in China. The US-headquartered multinational reported fulfilling the obligation by July (see 3M, 2020a; Bown, 2020d).

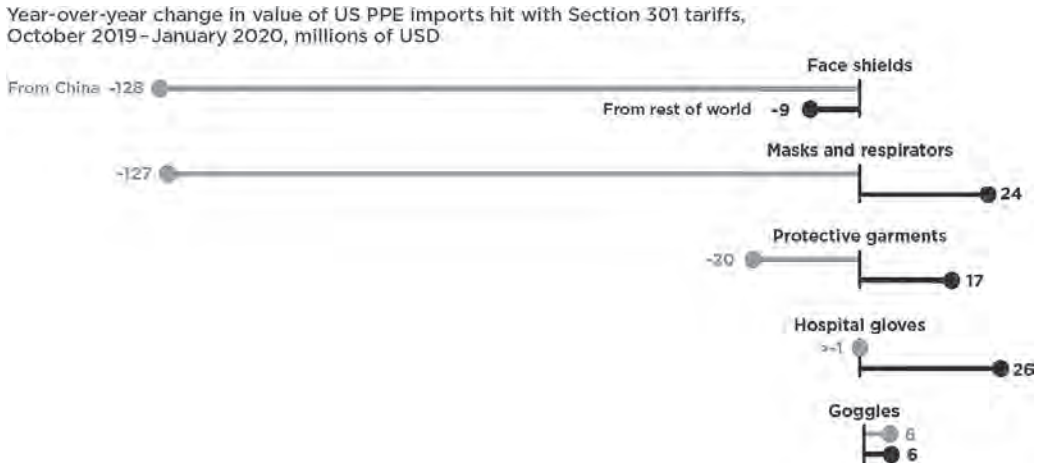


Figure 4 US disadvantage in pandemic preparedness due to trade war with China.
Source: US Census.
Note: personal protective equipment products included in US Section 301 List 4A and subject to new 15% tariff in September 2019.

Figure 5 shows respirator exports to the USA from Shanghai, the location of a 3M respirator plant in China. It plots, by Chinese province, the change in Chinese export volumes (Figure 5a) and the change in Chinese export prices (Figure 5b) for shipments to the USA relative to the rest of the world over the period of that arrangement. Compared to the average across provinces, Shanghai had slightly higher export volume increases, and slightly lower export price increases, to the USA relative to the rest of the world. This is consistent with meeting the DPA objective. However, it is possible that 3M would have increased imports from its Shanghai plant to the USA anyway, or that alternative Chinese suppliers in other provinces exported less to the USA to make up for those 3M orders.

5.3 The Defense Production Act's export controls, and PPE sales to Canada and Mexico

In the face of PPE shortages, a second extraordinary element of the April DPA invocation was US imposition of export controls on respirators, masks, and hospital gloves. On April 3, 3M (2020b) released a surprising statement that the US administration had asked it to "cease exporting respirators that we currently manufacture in the United States to the Canadian and Latin American markets" even though there would be "significant humanitarian implications of ceasing respirator supplies to healthcare workers in Canada and Latin America, where we are a critical supplier of respirators." The initial version of the regulation ignored the concern and limited US exports to Canada and Mexico; the restriction was only removed in the revised version published on April 17.[15]

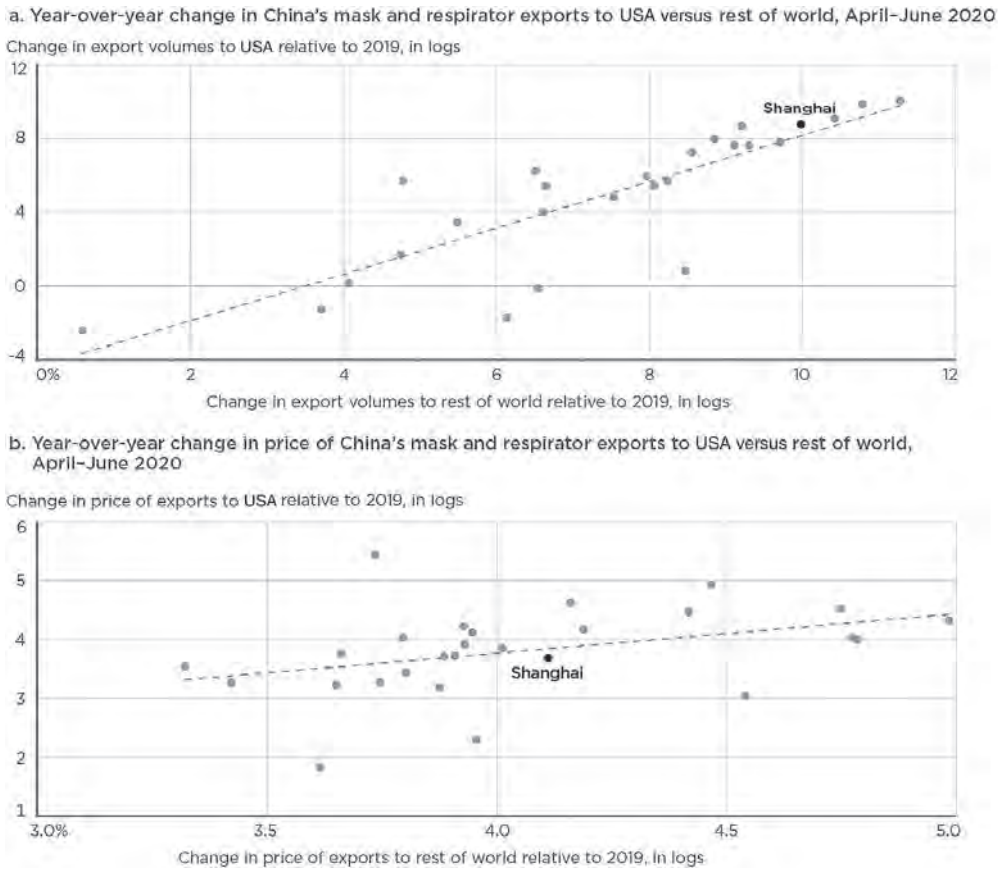


Figure 5 Effect of US arrangement with 3M on export volume and price of masks and respirators from Shanghai.
Source: Chinese Customs.
Note: Dots show data for 22 Chinese provinces, 5 autonomous regions, and 4 municipalities. Not shown are data for Taiwan and the special administrative regions of Hong Kong and Macau. Masks and respirators defined as HS code 6307909889. Data for each panel defined as log[(April + May + June 2020) − (April + May + June 2019)].

Nevertheless, the trade data alone provide little evidence that DPA negatively impacted US exports (Figure 6).[16] Canada and Mexico dominate US exports for each product, with the exception of air-purifying respirators, and export volumes ended up higher in 2020 than in 2019. US export volumes to Canada and Mexico of respirators and masks, for example, were 26% higher in the last three quarters of 2020 relative to 2019. US export prices to Canada and Mexico peaked in April 2020 at 120% higher than pre-pandemic levels, before declining over the rest of 2020. Again, this was much less than the Chinese export price increase, raising the question of whether a side effect of US export controls was to limit PPE price increases during extreme global scarcity. Air-purifying respirators – a product not previously discussed – are even more sophisticated than an N-95 respirator. After a US export surge in March, foreign sales fell

alongside the imposition of export controls in April, though they increased again later in 2020.

While US exports of these products were higher overall in 2020 relative to 2019, it remains unknown how much higher they would have been without the controls. Estiing th olicy's i c us ccount for the li ly incr s s in bo h do s ic (US) demand as well as foreign import demand; these would have competing effects on US export volumes, independent of the export control policy. The capacity expansion of the US industry (described next) would also increase export volumes, ceteris paribus.

Nevertheless, Canada responded by implementing industrial policy to reduce at least some of its future PPE import dependence on the USA. In August 2020, the governments of Canada and the province of Ontario announced subsidies for a 3M plant to manufacture N-95 respirators domestically (Ontario, 2020).

5.4 US industrial policy in support of the PPE supply chain in 2020

The US government eventually also deployed industrial policy, in the form of $1.2 billion of subsidies over the next year, to directly expand domestic PPE production capacity. It started by subsidizing domestic facilities producing N-95 respirators, beginning in mid-April 2020 (Table 2). Overall, it made nearly $800 million of publicly funded investments in American PPE production capacity expansion, as well as for

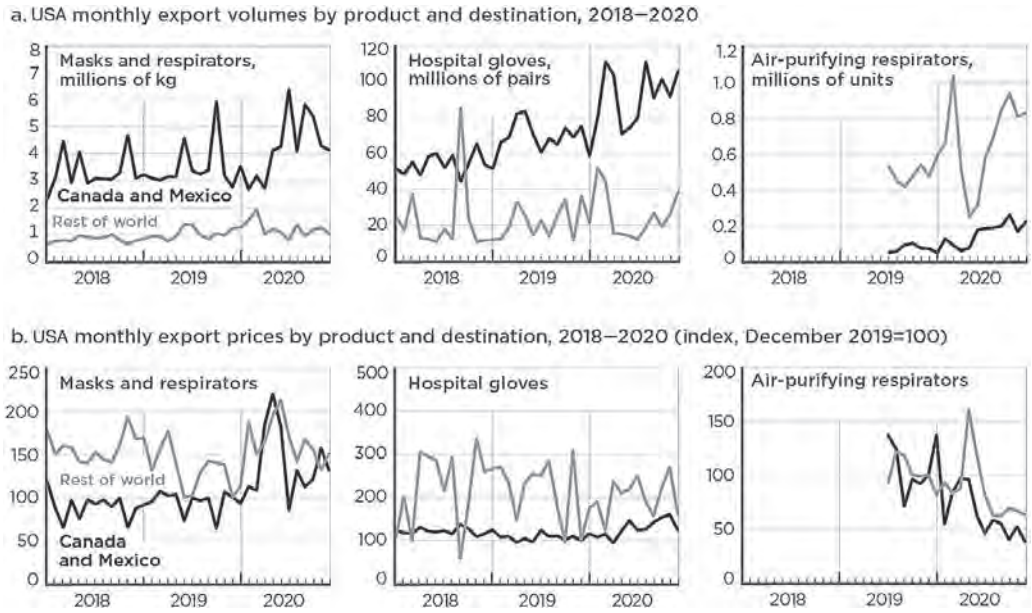


Figure 6 US personal protective equipment exports subject to export controls starting in April 2020.

Source: US Census.

Note: Export price in trade unit values. Data for air-purifying respirators start in July 2019.

inputs along the PPE supply chain, in 2020. Collaboration between the Department of Defense (DOD) and HHS, led by the DOD's Joint Acquisition Task Force and funded through the CARES Act, paid subsidies to 3M, O&M Halyard, Honeywell, Crosstex, and Medline Industries to add product lines for N-95 respirators or surgical masks. Fr man Manufactur ng and oth rs r c v d subs d s to sca up pr duct n of h sp - tal gowns. Funding was sent to Hollingsworth & Vose (filters), Lydall (meltblown filtration media), and NPS (meltblown fiber) to expand production of key inputs needed for PPE manufacturers of those surgical masks, respirators, and hospital gowns. Then, in May and June of 2021, the US government spent over $400 million on a half dozen companies to expand capacity – including for key raw material inputs – for nitrile glove production.

Late in 2020, the Trump administration also began a series of actions to potentially withdraw PPE from US commitments under the WTO's Government Procurement Agreement (GPA). Like import tariffs, this would force consumption of locally produced PPE – even if more costly relative to imports. After initially signaling potential support for the policy, the Biden administration reversed course in April 2021, possibly because it would have resulted in trading partner retaliation by withdrawing their own commitments under the GPA, hurting US exporters in other sectors.

In summary, the US policy actions to rebuild or expand the domestic PPE manufacturing sector were unlikely to have affected product availability until late 2020 at the earliest, with the exception of N-95 respirator capacity expansion. Nevertheless, the subsidization combined with the demand shock induced entry by American firms and changed the domestic industry landscape. However, a few months into 2021 prices had normalized sufficiently that some new entrants were starting to organize to request import protection. In May, for example, the American Mask Manufacturer's Association (2021), representing 26 small businesses, wrote to President Biden alleging that China was now "dumping masks on the US market at well below actual costs" and that if this practice continued "54% of our production will go offline in 60 days and 84.6% in less than a year."

US pandemic policy, as well as changing economic conditions, impacted industry structure in ways that also raised important questions for future preparedness policy.

6. Policy Implications

In the face of a global pandemic that created a surge in demand for PPE, an optimal policy mix for a major industrial economy should have involved three components: (i) Incentivize the domestic industry to add capacity and surge production as quickly as possible; and for the period during which surge capacity is ramping up and not yet available, rely on the combination of (ii) previously stockpiled PPE and (iii) imports. For the USA, COVID-19 revealed problems with all three parts of that strategy.

Table 2 US industrial policy for expanding the PPE manufacturing supply chain in 2020–2021

Date	Department of Defense policy action
2020	
April 11	$132.4 million combined to 3M ($76 million), O&M Halyard ($29 million), and Honeywell ($27.4 million) for N-95 respirator production expansion
May 6	Additional $126 million to 3M for N-95 respirator production expansion
May 28	$2.2 million to Hollingsworth & Vose for production expansion of N-95 ventilator filters and N-95 respirators
June 22	Memorandum of Understanding with US International Development Finance Corporation (DFC) to use $100 million of CARES Act funding to finance projects to help reshore production, including of PPE
June 19	$13.5 million to Lydall for meltblown filtration media production expansion
July 17	$3.5 million to Crosstex for surgical mask production expansion
July 24	$2.75 million to NPS for meltblown fiber line production expansion
September 11	$136 million to five companies for reusable isolation gown production expansion
September 14	$335 million to nine companies for disposable isolation gown production expansion
November 10	Additional $37 million to 3M for N-95 respirator production expansion
November 13	$6.18 million to Medline for surgical mask production expansion
November 20	$565,000 to Freeman Manufacturing for disposable gown production expansion
December 2	$2.5 million to Hollingsworth & Vose for filter media production expansion
2021[†]	
May 18	$56 million to Rhino Health for nitrile gloves production expansion
May 26	$13.1 million to Renco Corporation for nitrile gloves production expansion
May 27	$63.6 million to US Medical Glove Company for nitrile gloves production expansion
May 28	Additional $35 million to Renco Corporation for nitrile gloves production expansion
May 28	$123.1 million to Blue Star NBR for nitrile butadiene rubber production expansion, a key raw material in nitrile gloves
June 17	$37.6 million to Showa Best Glove for nitrile gloves production expansion
June 21	$96.1 million to United Safety Technology for nitrile gloves production expansion

Source: Constructed by the author from Department of Defense announcements and other sources.
[†]Through June 30, 2021.
PPE, personal protective equipment.

6.1 Production problems and policy

Why did the US government's industrial policy response take so long?[17] US government subsidies for PPE capacity expansion began to roll out only in April 2020; most

were not announced until the second half of the year, and some not until mid-2021. Yet private companies saw the changing conditions earlier – 3M, for example, announced capacity expansions beginning in January 2020. Even by February, other parts of the US government recognized PPE shortages – see, for example, the Congressional testimony of HHS Secretary Alex Azar (see CSPAN, 2020).

One explanation is that the government lacked basic information about domestic PPE production. Policymakers cannot target subsidies for PPE manufacturing companies that cannot be found. Missing information included how much and where domestic capacity existed prior to the pandemic, how quickly production could be expanded, and what resources (and other critical inputs in the supply chain) would be needed to make that happen.

Second, for some specific products, there may have been too little domestic production capacity altogether. "You can't surge zero," quipped John Polowczyk, in reference to America's apparent de minimis production of hospital gloves at the outset of the pandemic. This is plausible, but more data and analysis are needed to determine for what products that was such a constraint.

To better support policy going forward, the USA must collect and maintain up-to-date, detailed data on domestic production and capacity for PPE. The relevant industries will need to be subjected to periodic "war games" or "stress tests" to ensure that policy can incentivize a sufficiently quick expansion to surge capacity levels in a future emergency.

6.2 Stockpile problems and policy

Buyers, distributors, and governments collectively held too little inventory in reserve in case of emergency, as was made evident by the early pandemic depletion of the HHS Strategic National Stockpile. The further lack of inventory held by the private sector was, in part, likely the result of cost pressure.[18] A more robust system of preparedness may require regulators to ensure that hospitals, medical distributors, and states – in addition to the federal government – maintain more inventory. Because holding inventories is costly, and profit incentives pressure that part of the supply chain to become more lean, there is a role for regulation.

A separate question involves determining the socially optimal size of stockpiles to manage and for regulators to help oversee. That determination requires detailed projections on demand as well as information on the state of domestic production capacity (as discussed above) and how quickly it can be scaled up under differing pandemic scenarios. One scenario involves a health threat concentrated in the USA: imports would be available, but immediate domestic production might not. In other scenarios, only foreign supplies are unavailable, or both – or neither – sources are available. The global and rolling COVID-19 lockdowns over 2020 and 2021 highlight the importance of geographically diversified production within the USA as well as internationally. Relying solely on domestic production would be excessively risky, as would relying on imports primarily from one source.

6.3 Import problems and policy

Imports were a critical source of PPE during COVID-19, and should arguably remain an important component of future supply diversification. At the same time, although more data and detailed analysis are needed, imports may have contributed to multiple prob ems that emerged during the pandemic.

While PPE imports from China over the latter part of 2020 undoubtedly saved American lives, the lack of available imports in February and March likely cost lives. The problem might be characterized as a perfect storm of events. The pandemic arose in Hubei, the largest exporting province in the largest exporting country of the protective garments needed globally by hospital workers. The fact that those Chinese supplies were taken off the global market just when the rest of the world needed them shows that excessive concentration of production is a legitimate worry for American – and global – public health preparedness.

The USA and its trading partners must have a more diversified portfolio of foreign production for PPE. Achieving that objective may require new policy incentives – and forms of international coordination – if strong economic forces of agglomeration work to concentrate production geographically or in favor of the status quo.

Perhaps more so for the USA than other countries, international diversification must be a priority. Trade can be a tool for preparedness only if there is trust between the importer and the exporter – that is, confidence that when times get tough for health or economic reasons, trade lanes will remain open. There is now precious little trust between the USA and China as geopolitical tensions between the two countries remain elevated.

For certain products, imports over the years may have also contributed to insufficient domestic production to enable the government to surge capacity expansion during an emergency. A permanent policy intervention may be needed if optimal pandemic responsiveness requires a larger minimum domestic industry size than would be sustainable under normal market conditions and free trade, due to a positive externality. Policymakers will find tariffs attractive – and tariffs may emerge if better policies are not developed (a group like the American Mask Manufacturer's Association, for example, could petition bureaucrats to impose antidumping duties). However, while import protection does help stimulate domestic production, it also raises prices for consumers (e.g. in the health care system, which is already costly in the USA and many other countries). A more efficient policy to target an externality and achieve a sufficient minimal level of domestic production would be a subsidy.

Overall, the USA should ensure a diverse portfolio of imports of PPE for pandemic preparedness. Foreign sources of production must be transparent and, ideally, imports should come from countries with which the USA has a relationship of trust, to be sure that the source country will share supplies when times are challenging. The pandemic revealed that not many countries always fit the transparency and trust criteria, including the USA.

7. Conclusion and International Policy Cooperation

The PPE shortages and use of extraordinary trade and industrial policy during the COVID-19 pandemic revealed significant failures in preparedness. Trade played a mixed role.

The experience has triggered considerable rethinking of international cooperation for trade in such medical supplies. Indeed, the new US president's early 2021 meetings with leaders from Japan and the Group of seven countries led to joint statements and communiqués prioritizing PPE supply chain resilience as well as greater geographic diversification of production. Achieving those objectives will require different incentives and forms of international trade policy cooperation than were in place before the pandemic.

A new framework is also needed to define the proactive international policy coordination required at the first signs of the next emergency. A cooperative response of countries jointly and transparently triggering surge production capacity for PPE would do much to prevent a repeat of 2020 – waiting too long, followed by knee-jerk export restrictions.

Finally, the analysis here has focused on major economies with the capacity to push for and sustain domestic PPE industries. That will not be a feasible strategy for many smaller countries with markets that cannot achieve viable economies of scale. For them, challenges in trade and stockpile management will persist. Yet even these countries can learn important lessons from the US experience, including the need for visibility into trading partners' domestic production capacity and export product availability. Transparency is essential for any country seeking to maintain preparedness for public health emergencies.

Notes

1 "Remarks by President Biden at Signing of an Executive Order on Supply Chains," White House, February 24, 2021.

2 See, for example, USITC (2020), Baldwin and Evenett (2020), Evenett (2020), Evenett *et al.* (2021), Espitia *et al.* (2020), Gereffi (2020), Hoekman *et al.* (2020), Leibovici and Santacreu (2020), and Miroudot (2020).

3 This PPE characterization is a by-product of trade statistics classification prior to the pandemic. For example, while surgical masks and N-95 respirators are different products, they were inseparable because they fell into the same code. A similar explanation holds for different types of protective garments. As one policy response, the USA created new product codes for N-95 respirators, surgical masks, and face shields in July 2020 and for surgical gowns in January 2021. Supporting Information Table S2 provides precise product classifications.

4 The interview with Polowczyk is in Bown and Keynes (2021).

5 Depending on the product, applied MFN tariffs ranged from 0% to 7% for the USA and 1.7% to 12% for the EU.

6 At this point in the pandemic, the US government and European Commission were shipping PPE to China for humanitarian purposes (Lenarčič, 2020; Pompeo, 2020).

7 The Lunar New Year means January and February data for China are notoriously challenging to seasonally adjust. Year-over-year comparison here cumulates January–March 2020 and the same months in 2019. China's initial data release of March 25, 2020 did not include separate data for January and February (Bown, 2020a,e).

8 Hubei was a much smaller export supplier of the other PPE products. In 2019, it was the source of 4% of China's exports of masks and respirators, and less for the other. For

diplomacy-related explanations of regional differences in China's mask and respirator exports through March 2020, see Fuchs *et al.* (2020).

9 See Lin (2020) and Chinese Customs (2020) for export quality controls, and Stevenson and May (2020) and Lau (2020) for quality concerns.

10 China's exports had been negatively impacted by prior failures to regulate product quality. Bai et al. (2021) document the export impact of the 2008 scandal involving melamine-contaminated infant formula.

11 Supporting Information Table S3 provides more detail on EU import prices and alternative sources of imports.

12 Supporting Information Table S4 provides more detail on US import prices and alternative sources of imports.

13 See the testimony in the Section 301 hearings cited in Bown (2020c). Bown (2020d) documents even later requests for tariff exclusions on pandemic-related imports. Unlike other countries, the USA never had a public discussion over suspending MFN tariffs on PPE.

14 Even with the Phase 1 agreement, the explicit policy goal was to keep tariffs on China in place to reduce the bilateral trade deficit (Bown, 2021).

15 The export controls were extended in August until December 2020 and then again until June 2021. Some products were later added and others subtracted. See Supporting Information Table S1.

16 In 2019, US exports of these products were $553 million to Canada and Mexico and $529 million to the rest of the world. In 2020, US exports were $874 million to Canada and Mexico and $643 million to the rest of the world.

17 This section draws in part from interviews with John Polowczyk (Bown & Keynes, 2021).

18 A separate problem involves how the limited PPE stockpiles (and federal acquisitions) were allocated within the USA – that is, according to emerging public health demands (i.e. "hotpots") versus some other formula, such as the share of the national population. See House Committee on Oversight and Reform (2020) and HHS (2020) for the Strategic National Stockpile.

References

3M (2020a). *Explore the Steps We Have Taken to Support the World's Response to the COVID-19 Pandemic*. Accessed 30 June 2021. Available from URL: https://www.3m.com/3M/en_US/company-us/coronavirus/

3M (2020b). 3M response to Defense Production Act Order. Press release, 3 April.

Abaluck J., Chevalier J., Christakis N.A., Forman H., Kaplan E.H., Ko A. & Vermund S.H. (2020). The case for universal cloth mask adoption and policies to increase supply of medical masks for health workers. Yale University Working Paper, 1 April.

American Mask Manufacturer's Association (2021). The majority of U.S. mask manufacturing will go offline in 60 days; 2,647 jobs already lost. Letter to President Joe Biden, 11 May.

Bai J., Gazze L. & Wang Y. (2021). Collective reputation in trade: Evidence from the Chinese dairy industry. *Review of Economics and Statistics* (forthcoming).

Baldwin R. & Evenett S. (eds.) (2020). *COVID-19 and Trade Policy: Why Turning Inward Won't Work*. London: CEPR Press.

Bown C.P. (2020a). COVID-19: China's exports of medical supplies provide a ray of hope. PIIE Trade and Investment Policy Watch, 26 March.
Bown C.P. (2020b). EU limits on medical gear exports put poor countries and Europeans at risk. PIIE Trade and Investment Policy Watch, 19 March.
Bown C.P. (2020c). Trump's trade policy is hampering the US fight against COVID-19. PIIE Trade and Investment Policy Watch, 13 March.
Bown C.P. (2020d). COVID-19: Trump's curbs on exports of medical gear put Americans and others at risk. PIIE Trade and Investment Policy Watch, 9 April.
Bown C.P. (2020e). China should export more medical gear to battle COVID-19. PIIE Trade and Investment Policy Watch, 5 May.
Bown C.P. (2021). The US-China trade war and phase one agreement. *Journal of Policy Modeling*, **43** (4), 805–843.
Bown C.P. & Keynes S. (2021). How America responded to its PPE shortage. *Trade Talks*, Podcast episode (forthcoming).
Chinese Customs (2020). Announcement no. 53 (2020) of the General Administration of Customs. 10 April.
CSPAN (2020). Health and Human Services Fiscal Year 2021 Budget request. Washington, 25 February.
Espitia A., Rocha N. & Ruta M. (2020). A pandemic trade deal: Trade and policy cooperation on medical goods. In: Baldwin R.E. & Evenett S.J. (eds.), *Revitalising Multilateralism: Pragmatic Ideas for the New WTO Director-General*. London: CEPR Press, 189–201.
European Commission (2020a). Commission implementing regulation (EU) 2020/402 of 14 March 2020 making the exportation of certain products subject to the production of an export authorization. *Official Journal of the European Union*, L 077I, 15 March. Brussels.
European Commission (2020b). Commission publishes guidance on export requirements for personal protective equipment, 20 March. Brussels.
Evenett S. (2020). Chinese whispers: COVID-19, global supply chains in essential goods, and public policy. *Journal of International Business Policy*, **3** (3), 408–429.
Evenett S., Fiorini M., Fritz J. *et al.* (2021). Trade policy responses to the COVID-19 pandemic crisis: Evidence from a new dataset. *The World Economy* (forthcoming).
Fuchs A., Kaplan L., Kis-Katos K., Schmidt S.S., Turbanisch F. & Wang F. (2020). Mask wars: China's exports of medical goods in times of COVID-19. *Covid Economics*, **42**, 26–64.
Gereffi G. (2020). What does the COVID-19 pandemic teach us about global value chains? The case of medical supplies. *Journal of International Business Policy*, **3** (3), 287–301.
Hoekman B., Fiorini M. & Yildirim A. (2020). COVID-19: Export controls and international cooperation. In: Baldwin R.E. & Evenett S.J. (eds.), *COVID-19 and Trade Policy: Why Turning Inward Won't Work*. London: CEPR Press, 77–87.
Honeywell (2020). Honeywell expands face mask production in Europe with new manufacturing line in the United Kingdom. Press release, 13 May.
House Committee on Oversight and Reform (HCOR) (2020). New document shows inadequate distribution of personal protective equipment and critical medical supplies to states. Press release, 8 April, Washington.
Hui W. (2020). China denies banning export of face masks. *CGTN*, 5 March.
Keynes S. (2020). New trade barriers could hamper the supply of masks and medicines. *The Economist*, 11 March.

Lau S. (2020). Netherlands recalls 600,000 face masks from China due to low quality. *South China Morning Post*, 29 March.
Leibovici F. & Santacreu A.M. (2020). International trade of essential goods during a pandemic. Working Paper no 2020-010, May, Federal Reserve Bank of St. Louis.
Lenarčič J. (2020). Statement by commissioner for crisis management on EU support to China for the Coronavirus outbreak. European Commission.
Lin L. (2020). China tightens customs checks for medical equipment exports. *Wall Street Journal*, 10 April.
Miroudot S. (2020). Resilience versus robustness in global value chains: Some policy implications. In: Baldwin R.E. & Evenett S.J. (eds.), *COVID-19 and Trade Policy: Why Turning Inward Won't Work*. London: CEPR Press, 117–130.
O'Keeffe K., Lin L. & Xiao E. (2020). China's export restrictions strand medical goods US needs to fight coronavirus, State Department says. *Wall Street Journal*, 16 April.
Ontario (2020). Ontario partners with Federal Government and 3M Canada on new N-95 respirator manufacturing facility. News release, Office of the Premier, 21 August.
Pompeo M.R. (2020). The United States announces assistance to combat the Novel Coronavirus. Press statement. *Secretary of State*, 7 February.
State Council. (2020). The State Council Information Office held a press conference on the progress of the prevention, control and treatment of the novel coronavirus epidemic. 6 March. Beijing.
Stevenson A. & May T. (2020). China pushes to churn out Coronavirus gear, but struggles to police it. *New York Times*, 27 March.
United States International Trade Commission (USITC) (2020). COVID-19 related goods: The U.S. industry, market, trade, and supply chain challenges. Publication Number: 5145, Investigation Number: 332–580, December. Washington.
US Department of Health and Human Services (HHS) (2020). Strategic National Stockpile, Washington.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1: Supporting Information
Table S1 Timeline of key COVID-19-related events for personal protective equipment (PPE), 2018–21
Table S2: Personal protective equipment and HS product codes.
Table S3 EU imports of PPE in 2020, by product and source

Exhibit

30

High-Quality Masks Reduce COVID-19 Infections and Deaths in the US

Running Head: Masks Reduce COVID-19 Infections and Deaths in the US

Authors: Erik Rosenstrom, BS[1], Buse Eylul Oruc, MS[2], Nathaniel Hupert, MD MPH[3], Julie Ivy, PhD[1], Pinar Keskinocak, PhD[2,4], Maria E. Mayorga, PhD[1], Julie L. Swann, PhD[1,5]

1. Department of Industrial and Systems Engineering, North Carolina State University, Raleigh, NC, United States of America

2. School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, United States of America

3. Weill Cornell Medicine Graduate School of Medical Sciences, Cornell University, New York, NY, United States of America

4. Rollins School of Public Health, Emory University, Atlanta, GA, United States of America

5. Biomedical Engineering, University of North Carolina, Chapel Hill, NC, United States of America

Correspondence: Julie Swann, jlswann@ncsu.edu; Office phone (919-515-6423); NC State University; 111 Lampe Drive; Raleigh, NC 27695

Manuscript Word Count: 2645

Declarations

Funding: Financial support for this study was provided in part in part by the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health, through Grant Award Number UL1TR002489. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The research was also supported by the Council of State and Territorial Epidemiologists and the Centers for Disease Control and Prevention, NC State University, the Fitts Department of Industrial and Systems Engineering at NC State, and the Georgia Institute of Technology. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. Several of the authors are employed by the state universities associated

with the sponsorship. We would like to acknowledge Dr. Nicoleta Serban of Georgia Tech for her work on a previous Covid-19 simulation and Dr. Mehul Patel for his leadership on the CovSim initiative. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. All of the authors are employed by the state universities associated with the sponsorship.

Conflicts of Interest/Competing Interests: There are no conflicts to declare.

Availability of data and material: Data used in the simulation is publicly available, and parameters are included with the supplemental document.

Code availability: Computer code is not presently available publicly as it is associated with the dissertation work of several PhD students.

Authors' contributions: All authors contributed to the analysis in the manuscript and the writing of the paper.

ACKNOWLEDGEMENTS

The project described was supported in part by the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health, through Grant Award Number UL1TR002489. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The research was also supported by the Council of State and Territorial Epidemiologists and the Centers for Disease Control and Prevention, NC State University, the Fitts Department of Industrial and Systems Engineering at NC State, the Georgia Institute of Technology, and the Cornell Institute for Disease and Disaster Preparedness. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. Several of the authors are employed by the state universities associated with the sponsorship. We would like to acknowledge Dr. Nicoleta Serban of Georgia Tech for her work on a previous Covid-19 simulation and Dr. Mehul Patel for his leadership on the CovSim initiative.

Title: High-Quality Masks Reduce COVID-19 Infections and Deaths in the US

ABSTRACT

Objectives: To evaluate the effectiveness of widespread adoption of masks or face coverings to reduce community transmission of the SARS-CoV-2 virus that causes COVID-19.

Methods: We created an agent-based stochastic network simulation using a variant of the standard SEIR dynamic infectious disease model. We considered a mask order that was initiated 3.5 months after the first confirmed COVID-19 case. We varied the likelihood of individuals wearing masks from 0-100% in steps of 20% (mask adherence) and considered 25% to 90% mask-related reduction in viral transmission (mask efficacy). Sensitivity analyses assessed early (by week 13) versus late (by week 42) adoption of masks and geographic differences in adherence (highest in urban and lowest in rural areas).

Results: Introduction of mask use with 50% efficacy worn by 50% of individuals reduces the cumulative infection attack rate (IAR) by 27%, the peak prevalence by 49%, and population-wide mortality by 29%. If 90% of individuals wear 50% efficacious masks, this decreases IAR by 54%, peak prevalence by 75%, and population-wide mortality by 55%; similar improvements hold if 70% of individuals wear 75% efficacious masks. Late adoption reduces IAR and deaths by 18% or more compared to no adoption. Lower adoption in rural areas than urban would lead to rural areas having the highest IAR.

Conclusions: Even after community transmission of SARS-CoV-2 has been established, adoption of mask-wearing by a majority of community-dwelling individuals can meaningfully reduce the number and outcome of COVID-19 infections over and above physical distancing interventions.

Highlights:

- This paper shows the impact of widespread adoption of masks in response to the COVID-19 pandemic, with varying levels of population adherence, mask efficacy, and timing of mask adoption.
- The paper's findings help inform messaging to policymakers at the state or local level considering adding or keeping mask mandates, and to communities to promote widespread adoption of high-quality masks.
- Adoption of masks by at least half of the population can reduce cumulative infections and population deaths by more than 25%, while decreasing peak prevalence by about 50%. Even greater marginal improvements arise with adoption rates above 70%. The benefits of adopting high-quality masks is above that achieved by mobility changes and distancing alone.
- Rural and suburban areas are at higher relative risk than urban areas, due to less distancing and lower adoption of masks.

INTRODUCTION

Since its introduction in early 2020, the SARS-CoV-2 coronavirus that causes COVID-19 has been widely circulating throughout the United States [1]. While there is growing empirical evidence that masks can be effective at reducing SARS-CoV-2 transmission via droplets [2, 3], the impact of population-based mask use on the COVID-19 pandemic is still poorly defined. Some states were early mask promoters (e.g., North Carolina or NC in June), some were late (e.g., North Dakota in November), and some have remained silent (e.g., South Dakota). This study prospectively assesses the effectiveness of face coverings at the state level and across urban, suburban, and rural counties under different mask efficacy and population adherence levels while providing additional evidence for the benefit of existing adoptions. We focus on masks because they are relatively inexpensive and logistically non-disruptive (compared to, e.g., physical distancing measures). Masks are also a fast intervention that can be adopted if SARS-Cov-2 mutates or if other viruses arise. We include analysis of state-wide effects along with those on subpopulations stratified by geography and demographics.

METHODS

We employ an agent-based stochastic network model with an SEIR framework [4] for the progression of SARS-CoV-2 [5, 6]. As in Keskinocak et al [6], we simulate interactions among agents, where transmission can occur daily in households, workplaces and schools, and community settings, with day/night differentiation in interactions. The model is also similar to Patel et al [7], where we study vaccine distribution. The network of agents is built using US Census data [8] at the tract level using household size, presence of children, and age groups (defined as <= 4,5- 9, 10-19, 20-64, or 65 years old and above). Model parameters include the reproductive rate (2.4 without interventions), hospitalization and mortality rates (by age group), an effective overall Infection Fatality Rate just under 0.5%, and asymptomatic and symptomatic transmission coefficients [5-7]. We also incorporate race/ethnicity in the model at the household level, as well as the existence of a COVID-19-associated health condition (diabetes) stratified by race/ethnicity using available data [9]. Full details are outlined in Supplementary Materials available online.

Analysis was performed for the representative state of North Carolina (NC), which has had SARS-CoV-2 transmission rate that puts it in the lower half of US states as quantified by

deaths per capita [1]. The population of 10.5 million people is represented with a sample of 1,017,720 agents. The simulation is seeded (day 1) with cumulative cases as of March 24, 2020 [1], where the initial cases are multiplied by 10 to account for underreporting [10] and scaled to the number of agents in the simulation; the simulation is run for 365 days. County cases are distributed to households randomly across each county's census tracts proportional to the tract population. We categorize each census tract by urban/suburban/rural status, where urban corresponds to Rural-Urban-Commuting-Area (RUCA) codes of 1,2 for urban; 6,7,8,9,10 for rural; and the remainder for suburban [11].

The model captures the likelihood an adult agent stays home over time using SafeGraph data [12] (see supplement for details) aggregated by month and census tract. SafeGraph data captures the presence of devices in homes or other settings over time and across census blocks, which we aggregated into tracts grouped by urbanicity (urban, rural, suburban) and median household income (4 quartiles statewide). An adult will work from home on a given day according to a probability drawn from their census tract's rate. For interactions with the community by age group, we assume the rate follows the same pattern as the workplace mobility data but with a smaller reduction compared to workplace (e.g., a 40% reduction in work attendance is associated with a 12% reduction in community interaction), see Supplemental Appendix). This is consistent with our comparisons of workplace mobility with other types of mobility [12, 13]. We assume mobility rates stabilize at month six levels. We do not assume a link between mobility and population infections as we did not see it consistently across locations when comparing mobility data and infections. The mobility data captures the fact that many people stayed home shortly after cases began rising (consistent with shelter-at-home orders given in NC and many other states) and mobility continues to be lower than pre-pandemic rates. We assume that households with a symptomatic Covid-19 infection will voluntarily quarantine (VQ), in line with the low quarantining rates from [6] (see Supplemental Appendix). Schools are virtual or closed initially, opening on month six with students rotating every other day. Anyone who is symptomatic stays home from school and away from work peer groups.

Unlike Keskinocak et al [6], a proportion of the population wears masks under early and late scenarios. In the early scenario, the adherence rate increases approximately monthly (days 6-94, where the last date corresponds to the state mask order in NC), linearly from 0 to the final adherence probability of (0, 40, 60, 80, or 100%). The "late" scenario, where mask adoption

reaches the highest level when the cumulative infections are approximately equal to 15% in the baseline cases (specifically, on day 282), corresponds to later mask orders such as the one issued in North Dakota [14]. In the baseline scenarios, we assume the mask adherence is homogeneous across the population, as supported by large-scale random sampling of the population from July 2020 [15]. In sensitivity analysis, we allow mask adherence to vary by urbanicity, [85%,75%,65%] or [85%,70%,55%] based on the November 2020 surveys conducted by Facebook [16], with values representing typical ones for states with higher or lower mask usage (e.g., NC that had early adoption vs ND that had late adoption). Based on recent experimental analysis [3], in the baseline scenarios we assume that masks reduce the infectivity to others and susceptibility by 50% each (mask efficacy); if two agents encounter each other in the simulation, the potential for encounter-based transmission reduction is multiplicative if both are wearing masks. We compare the baseline mask efficacy with scenarios where higher quality, more efficacious masks are employed (e.g., 80+% reduction in transmission and susceptibility risk, as may be achieved by surgical, N95, or even multi-layer sewn masks). Note that the population-level effectiveness of masks can be computed as adherence multiplied by efficacy, e.g., 75% adherence times 50% efficacy equals 37.5% effectiveness. For our no-intervention control, we assume there are no interventions throughout the pandemic and mobility is as normal; there is also a scenario with 0% mask adherence that has changes in mobility. Note that we do not assume a direct relationship between rising infections and other aspects such as distancing (captured by mobility), mask usage, or changes in fatality rates.

The model is validated against reported hospitalizations and deaths in NC as of 11/1/2020, where the validation accounts for the fact that not all positive cases are lab-reported (See the Supplemental Appendix). All simulation output values are adjusted to the true population of 10.49 million.

We compute the infection attack rate (IAR) of all infections over the time horizon, the peak percentage of the population simultaneously infected, the peak count of hospitalizations, and the mortality of the total population (total deaths or a mortality rate for subpopulations. Since the model is stochastic, we quantify the mean and standard deviation values over 15 replications, where the number of replications balances reductions in variability of measures such as IAR and total deaths with computational time. We provide values at the state level, by county, and stratified by urban/suburban/rural status. In the studied state, 75% of people

are in urban areas, 15% in suburban, and 10% in rural areas. We use the phrase percentage points if referring to absolute differences or give a percentage if the change is relative (e.g., a drop from 30% to 20% IAR is 10 percentage points or a 33% reduction).

The funding supported the students performing the analysis; it had no relationship with the design of the experiments, interpretation of the results, or determining the conclusions.

This study uses publicly available de-identified data and does not require IRB approval.

RESULTS

Even at low levels of effectiveness, mask-wearing reduces cumulative infections, peak infections, hospitalizations, and mortality for COVID-19 (Table 1). If 75% of individuals wear 50% efficacious masks (population-level effectiveness of 37.5%), the IAR, peak infection, and deaths are reduced by 37%, 68%, and 47%, respectively, vs. no mask use, even in the setting of effective physical distancing measures. .

Higher mask adherence leads to improvement in each metric, and the improvement is not necessarily linear. Increasing adherence from 50% to 75% has a higher incremental improvement in IAR points (5.4) than increasing from 0 to 25% or 25% to 50% (3.6 and 4.3, respectively). A similar improvement is seen in the number of deaths, which drops by 25.8% when adherence increases from 50% to 75%.

Notably, in the best scenario we studied where 90% of people wear a mask that is 50% efficacious, this results in an almost 50% reduction in IAR to 13.7% (compared to 21.7% with 50% adherence); additionally, peak hospitalizations decrease by 51% and deaths by 36.7% in this scenario.

If higher quality masks are worn, then all metrics improve. As with mask adherence, the incremental improvement is nonlinear. For example, the IAR decreases incrementally by 7, 8.4, and 2.9 percentage points over the previous value as mask efficacy increases (from 25% to 50%, then to 75%, then to 90%, respectively, all with adherence of 70%). The incremental reduction in the number of deaths is also highest as efficacy increases from 50 to 75%.

Performance measures are better if mask adherence is adopted early (starting week 1), but late mask adherence (starting week 40) still improves population-based outcomes (see Figure 1, which shows the prevalence over time for mask adherence, efficacy, and timing, and Table

1 for corresponding summary values). Early adoption has a cumulative IAR that is 4.8 to 8.7 percentage points better than the corresponding late adoption case, depending on other parameters. However, even late adoption of mask use with adherence levels of 50%-90% improves the cumulative IAR over the no mask case by 3.1-7.2 percentage points, and reduces the number of deaths by 12.1%-25.6%. Late adoption with high mask adherence (90%) is marginally better than early adoption with low adherence (25% or 50%).

The results vary somewhat by RUCA county type, see Table 1 and Figure 2. Note that mobility differences across area types (shown in the Supplemental Appendix) indicate that urban areas have stayed home at higher rates than other areas, d with rural areas remaining the most mobile. In the baseline scenario (50% adherence): the IAR for suburban areas is about two points higher than that of urban and rural areas (23.59% versus 21.5% and 21.62%). The difference between urban and rural areas is biggest with 0% mask adherence (25.7% versus 26.3%) and smallest with 90% adherence (13.63% and 13.64%).When mask adherence differs across geographies as seems indicated in recent data, then the IAR tends to be highest for rural areas.

Examination of results across demographics indicate that Whites have the lowest IAR and Blacks have the highest mortality rate, findings that are consistent across all masking scenarios.

Finally, the impact of masks on simulation outcomes is additive to the effects of changes in mobility and voluntary quarantining. The control with no masks and no interventions had an IAR of 58%, peak prevalence of 4.7%, 49,000 hospitalizations, and more than 27,000 deaths, while the scenario with mobility changes but 0% mask adherence resulted in 29.6% IAR, 0.74% peak prevalence, 8390 hospitalizations, and fewer than 15000 deaths. In comparison, the results with 50% efficacy and 70% adherence or with 50% adherence and 75% efficacy provided additional improvement over no masks, with IAR below 18%, peak prevalence below 0.3%, hospitalizations below 3300, and deaths below 8600.

DISCUSSION AND CONCLUSIONS

According to this dynamic model of COVID-19 spread, widespread use of masks after community transmission has begun decreases the impact of the pandemic, even when the overall effectiveness of the intervention is less than 50%. This finding is consistent with empirical findings using publicly reported state-based data from April 1 to May 21, 2020 [17]

and a deterministic aggregate model of disease spread [18]. Our modeled scenarios suggest 40% or higher reductions in infections and mortality, over and above those resulting from physical distancing interventions in place, for scenarios with adherence 70% or greater, at least 50% efficacy, and early adoption. Even adoption of mask use many months later in the outbreak can lead to 18% or more reductions in IAR and deaths. We stress that since mask use does not reduce disease transmission to zero in any scenario, other non-pharmaceutical (and future pharmaceutical) interventions need to be maintained.

The finding that rural and suburban areas are at risk for high IAR is consistent with the recent spread of COVID-19 well beyond urban areas [1]. It is somewhat surprising that suburban areas are at higher risk in some scenarios, although these locales have relatively high population density compared to rural areas and their mobility changes have been less than in urban areas (see Supplemental Appendix). Differences in timing of mask adoption and in adherence across geographical areas may eventually be associated with much higher IAR in rural areas that adopted masks late compared to urban ones that adopted early, which was as much as 10 percentage points in the IAR difference in some scenarios from Figure 2. Our demographic results highlight continued disparities for Blacks even under mask mandates, and our results likely underestimate the true demographic gap in infections or mortality.

Our findings lend further support to states or other localities considering mask orders and evidence that such practices have been effective. Improving the quality of masks worn also has the potential to improve population health, e.g., by shifting people from coverings like neck gaiters or bandanas with lower efficacy to masks with multiple fabric layers or special filtration material like N95 masks. Public health organizations should consider adding this to messaging around face masks. In our discussions with health departments, the ability to quantify the improvement from increased mask adoption is also valuable in targeting messaging to the population.

Limitations of this modeling study include potential mischaracterization of SARS-CoV-2 transmission and illness-causing mechanisms, as well lingering data deficits regarding age-based hospitalizations and mortality rates. Some of the key assumptions that are critical include the efficacy of masks and the adherence in the population. The survey data used captures whether people report wearing masks “most or all of the time”, which is insufficient as a true measure of adherence. We do not account for bias in who is wearing a mask, in terms of their social networks or their mobility patterns. We do not incorporate behavioral

changes with respect to rising infections; our review of the recent mobility and infection data did not support this consistently across all areas and time periods (unlike earlier analysis by [19]). If this assumption is wrong then we overestimate the impact of masks. If crowded hospitals lead to an increased fatality rate then we underestimate the value of masks in reducing COVID-19-related mortality. The results do not account for vaccination, which we study in a companion paper [7].

In future research, it may be useful to examine the impact of potential covariance of of likelihood of vaccine uptake and likelihood of mask use, and consequent implications for health equity across diverse affected populations.

CONCLUSION

This simulation of the impact of mask use to counter ongoing COVID-19 pandemic outbreaksprovides evidence that transmission of the SARS-CoV-2 virus can be meaningfully reduced if many people wear masks, especially masks of higher quality. The effect is greater at higher levels of adherence and with masks of higher efficacy, but still meaningful at the lowest levels of each examined. Rural areas with low mask adherence are at particularly high risk for negative outcomes. Specific quantitative values of 20% benefit (late adoption) or 40% (early adoption) may help improve messaging.

Table 1: The results of the mask adherence and efficacy scenarios are shown for the metrics of IAR, peak prevalence rate, peak hospitalizations, and deaths for a state population of 10.5 million with mean (stdev) displayed. Highlighted rows are used in figure comparisons.

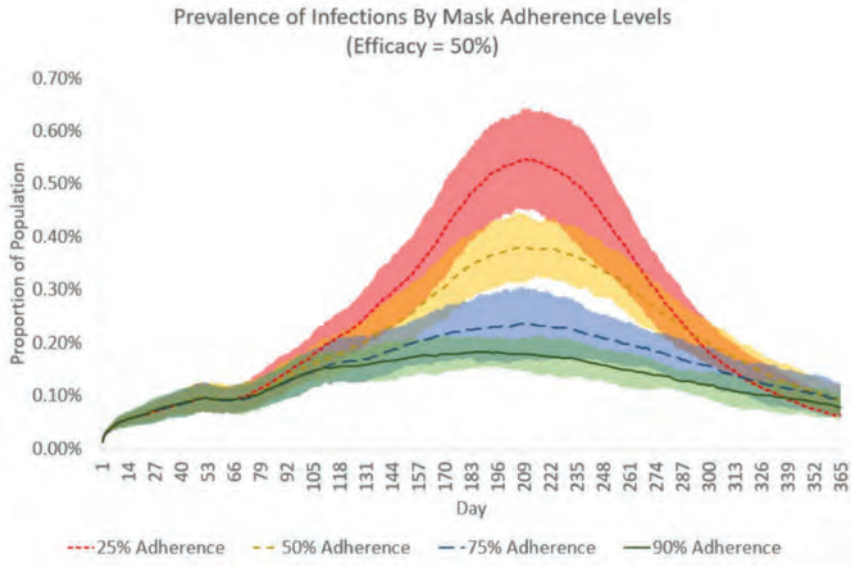
Mask Adherence	Mask Efficacy	IAR	Peak Prevalence Rate	Peak Hospital-izations	Deaths (in population of 10.5M)	N/A
Mask Adherence Experiments (Initial Shelter, Low Voluntary Quarantine throughout, School cancelled, early adoption of masks)						
0%, Overall	N/A	29.6% (0.20%)	0.74% (0.07%)	8,387 (834)	14,557 (528)	
25%	50%	26.0% (0.21%)	0.547% (0.058%)	6,195 (645)	12,316 (447)	
50%	50%	21.7% (0.22%)	0.38% (0.030%)	4,335 (418)	10,351 (544)	
75%	50%	16.3% (0.22%)	0.237% (0.035%)	2,741 (463)	7,681 (483)	
90%	50%	13.7% (0.22%)	0.184% (0.019%)	2,106 (280)	6,548 (544)	
Mask Efficacy Experiments (Initial Shelter, Low Voluntary Quarantine throughout, School cancelled, early adoption of masks)						
70%	25%	24.9% (0.21%)	0.494% (0.04%)	5,711 (537)	12,009 (480)	
70%	50%	17.9% (0.22%)	0.279% (0.030%)	3,257 (404)	8,630 (468)	
70%	75%	9.5% (0.20%)	0.117% (0.016%)	1303 (265)	4,461 (431)	
70%	90%	6.6% (0.18%)	0.094% (0.024%)	1,022 (187)	3,143 (241)	

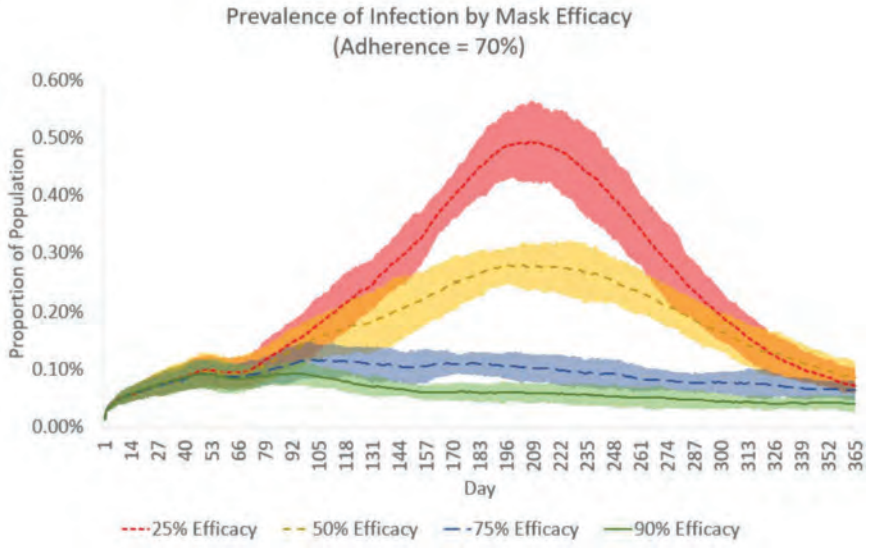
Delayed Mask Mandate (15% of the population has been infected)						
Mask Adherence	**Mask Efficacy**	**IAR**	**Peak Prevalence Rate**	**Peak Hospital-izations**	**Deaths (in population of 10.5M)**	
25%	50%	28.29% (0.2%)	0.66% (0.01%)	7,351 (707)	13,580 (697)	
50%	50%	26.46% (0.21%)	0.58% (0.014%)	6,566 (530)	12,793 (477)	
75%	50%	24.14% (0.21%)	0.49% (0.012%)	5,536 (497)	11,601 (547)	
90%	50%	22.44% (0.22%)	0.44% 0.021	5,019 (494)	10,827 (475)	
Control case with no interventions						
From the beginning: No masks, no mobility changes, yes school. Usual status.		58% (0.14%)	4.76% (0.2%)	49,142 (1,379)	27,982 (483)	
Geographical Analysis (Early Adoption)						
		Urban, IAR	Suburban, IAR	Rural, IAR	Overall, IAR	
Cases by March 24		0.003%	0.00025%	0.00027%	0.0023%	
0% masks (Initial shelter, Low VQ, schools cancelled)		29.07% (0.05%)	32.81% (0.14%)	30.03% (0.23%)	29.6% (0.20%)	
25%	50%	25.7% (0.31%)	28.1% (0.26%)	26.3% (0.4%)	26.0% (0.21%)	
50%	50%	21.5%	23.59%	21.62%	21.7%	

		(0.15%)	(0.3%)	(0.25%)	(0.22%)	
75%	50%	16.16% (0.54%)	17.82% (0.69%)	16.23% (0.26%)	16.3% (0.22%)	
90%	50%	13.63% (0.19%)	15.07% (0.37%)	13.64% (0.32%)	13.7% (0.22%)	

Demographic Analysis for IAR and Mortality Rate (MR)

Mask Adherence	**Mask Efficacy**	**White IAR**	**Black IAR**	**Hispanic IAR**	**Other IAR**	**White MR**	**Black MR**	**Hispanic MR**	**Other MR**
0% Masks		27.87% (0.22%)	31% (0.30%)	30.9% (0.38%)	31% (0.58%)	0.14% (0.00%)	0.15% (0.01%)	0.12% (0.01%)	0.1% (0.02%)
25%	50%	25.14% (0.22%)	27.66% (0.30%)	27.66% (0.38%)	27.56% (0.58%)	0.11% (0.00%)	0.13% (0.01%)	0.1% (0.01%)	0.08% (0.02%)
50%	50%	20.% (0.22%)	23.2% (0.31%)	23.1% (0.38%)	23.65% (0.58%)	0.09% (0.00%)	0.11% (0.01%)	0.09% (0.01%)	0.07% (0.02%)
75%	50%	15.5% (0.24%)	17.8% (0.32%)	17.7% (0.39%)	17.5% (0.59%)	0.07% (0.00%)	0.08% (0.01%)	0.07% (0.01%)	0.05% (0.01%)
90%	50%	13.1% (0.24%)	14.9% (0.33%)	15% (0.39%)	15.2% (0.59%)	0.06% (0.00%)	0.07% (0.01%)	0.06% (0.01%)	0.04% (0.01%)





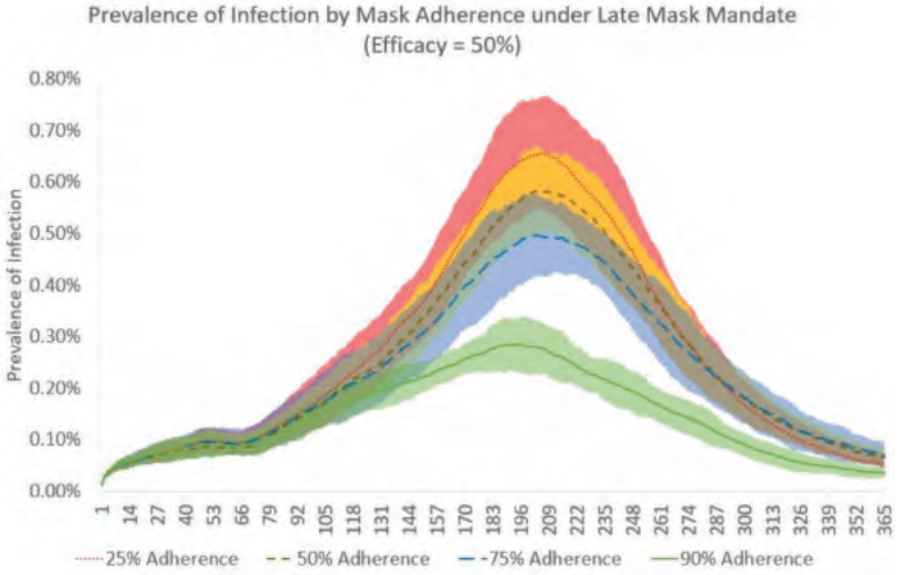


Figure 1: Prevalence of infectious people over time is shown. Subfigure (top) shows prevalence of cases with different levels of population adherence (25%, 50%, 75%, 90%) while mask efficacy is 50%. Subfigure (middle) shows prevalence with different levels of mask efficacy (25%, 50%, 75%, 90) while population adherence is 70%. Subfigure (bottom) shows prevalence with different levels of population adherence under late adoption of masks.

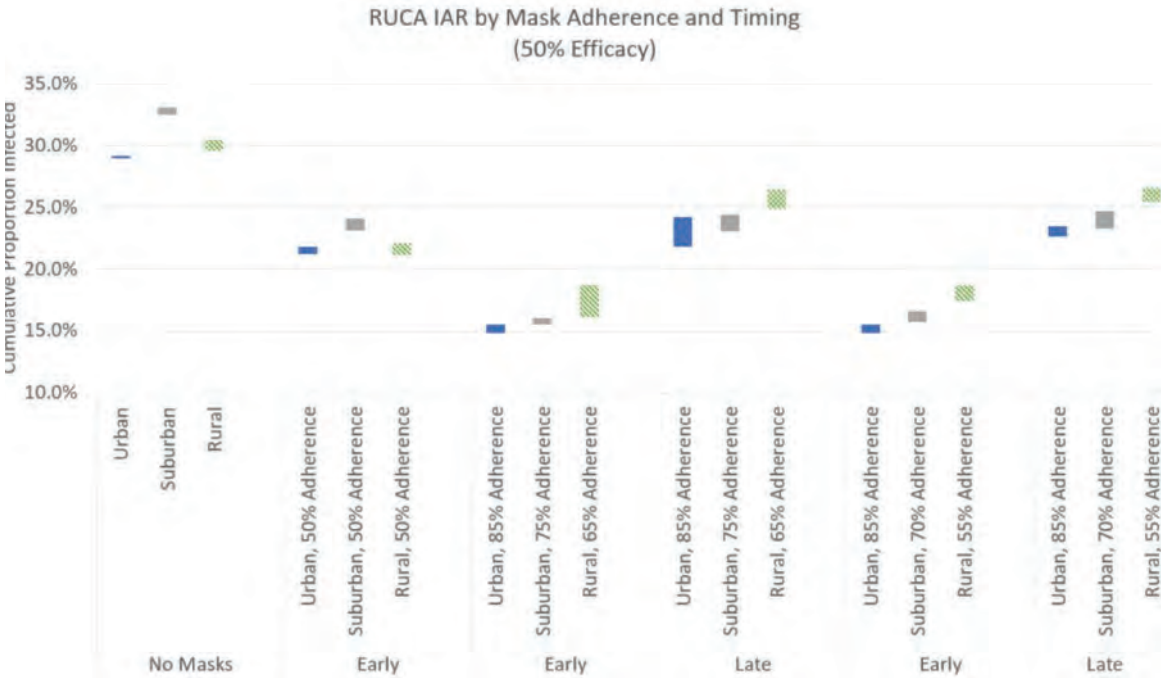


Figure 2: Infection Fatality Rates for areas categorized by urban, rural, and suburban status, where the band is ± 2 standard deviations around the daily mean for all areas in that category across all replications. Mask adherence is shown for 0% adherence, 50% across all areas with early adoption, and two scenarios of differing adherence levels by early and late timing.

REFERENCES

1. The New York Times. *Coronavirus in the U.S.: Latest Map and Case Count*. 2020 [cited 2020 18 Aug]; Available from: https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html.
2. Chu, D.K., et al., *Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis.* Lancet, 2020. **395**(10242): p. 1973-1987.
3. Fischer, E.P., et al., *Low-cost measurement of facemask efficacy for filtering expelled droplets during speech.* Science Advances, 2020.
4. Shi, P., et al., *Modelling seasonality and viral mutation to predict the course of an influenza pandemic.* Epidemiology and Infection, 2010. **138**(10): p. 1472-1481.

5. Keskinocak, P., et al., *Evaluating Scenarios for School Reopening under COVID19.* medRxiv, 2020: p. 2020.07.22.20160036.
6. Keskinocak, P., et al., *The impact of social distancing on COVID19 spread: State of Georgia case study.* Plos One, 2020. **15**(10).
7. Patel, M.D., et al., *The Joint Impact of COVID-19 Vaccination and Non-Pharmaceutical Interventions on Infections, Hospitalizations, and Mortality: An Agent-Based Simulation.* MedRXiv, 2020.
8. U.S. Census Bureau. 2018; Available from: data.census.gov.
9. America's Health Rankings. *Annual Report*. 2019 [cited 2020 12/30/2020]; Estimates of Diabetes by state from CDC's BRFSS]. Available from: https://www.americashealthrankings.org/explore/annual/measure/Diabetes/state/NC.
10. Havers, F.P., et al., *Seroprevalence of Antibodies to SARS-CoV-2 in 10 Sites in the United States, March 23-May 12, 2020.* Jama Internal Medicine, 2020. **180**(12): p. 1576-1586.
11. US Department of Agriculture. *Rural-Urban Commuting Area Codes*. 2010 [cited 2020 July 30]; Available from: https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes.aspx.
12. Safegraph. Available from: http://safegraph.com/.
13. Google. *COVID-19 Community Mobility Reports*. 2020 [cited 2020 12/29]; Available from: https://www.google.com/covid19/mobility/.
14. AARP. *State-by-State Guide to Face Mask Requirements*. 2020 12/24/2020 [cited 2020 12/29]; Available from: https://www.aarp.org/health/healthy-living/info-2020/states-mask-mandates-coronavirus.html.
15. The New York Times. *A Detailed Map of Who Is Wearing Masks in the U.S.* 2020 [cited 2020 18 Aug]; Available from: https://www.nytimes.com/interactive/2020/07/17/upshot/coronavirus-face-mask-map.html.
16. Facebook. *Percentage of people who report wearing a mask most or all of the time while in public, based on surveys of Facebook users*. 2020 12/26/2020 [cited 2020 12/1]; Available from: https://delphi.cmu.edu/covidcast/?date=20201226&region=42003&sensor=fb-survey-smoothed_wearing_mask.
17. Lyu, W. and G.L. Wehby, *Community Use Of Face Masks And COVID-19: Evidence From A Natural Experiment Of State Mandates In The US.* Health Affairs, 2020. **39**(8): p. 1419-1425.
18. Eikenberry, S.E., et al., *To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic.* Infectious Disease Modelling, 2020. **5**: p. 293-308.
19. Goolsbee, A. and C. Syverson, *Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020.* Journal of public economics, 2021. **193**: p. 104311-104311.

Exhibit

31

PANDEMIC TECHNOLOGY PROJECT

Inside the risky bat-virus engineering that links America to Wuhan

China emulated US techniques to construct novel coronaviruses in unsafe conditions.

By Rowan Jacobsen

June 29, 2021

MS TECH | AP

In 2013, the American virologist Ralph Baric approached Zhengli Shi at a meeting. Baric was a top expert in coronaviruses, with hundreds of papers to his credit, and Shi, along with her team at the Wuhan Institute of Virology, had been discovering them by the fistful in bat caves.

In one sample of bat guano, Shi had detected the genome of a new virus, called SHC014, that was one of the two closest relatives to the original SARS virus, but her team had not been able to culture it in the lab.

Baric had developed a way around that problem—a technique for "reverse genetics" in coronaviruses. Not only did it allow him to bring an actual virus to life from its genetic code, but he could mix and match parts of multiple viruses. He wanted to take the "spike" gene from SHC014 and move it into a genetic copy of the SARS virus he already had in his lab. The spike molecule is what lets a coronavirus open a cell and get inside it. The resulting chimera would demonstrate whether the spike of SHC014 would attach to human cells.

If it could, then it could help him with his long-term project of developing universal drugs and vaccines against the full spectrum of SARS-like viruses that he increasingly considered sources of potential pandemics. A SARS vaccine had been developed, but it wasn't expected to be very effective against related coronaviruses, just as flu shots rarely work against new strains. To develop a universal vaccine that will elicit an antibody response against a gamut of SARS-like viruses, you need to show the immune system a cocktail of spikes. SHC014 could be one of them.

If you study a hundred different bat viruses, your luck may run out.

-Ralph Baric, University of North Carolina

Baric asked Shi if he could have the genetic data for SHC014. "She was gracious enough to send us those sequences almost immediately," he says. His team introduced the virus modified with that code into mice and into a petri dish of human airway cells. Sure enough, the chimera exhibited "robust replication" in the human cells—evidence that nature was full of coronaviruses ready to leap directly to people.

While Baric's study was in progress, the National Institutes of Health announced that it would temporarily halt funding for "gain of function" research—experiments that make already dangerous viruses more virulent or transmissible—on SARS, MERS (which is also caused by a coronavirus), and influenza until the safety of such research could be assessed. The announcement brought Baric's work to a standstill.

Related Story

Baric was a legend in the field, but no matter how many safety precautions are taken, there is always a chance that a never-before-seen virus can escape and trigger an outbreak. Baric felt that the extreme measures he took in the lab minimized the risk, and in fact made his work categorically different from the high-risk influenza work the NIH had been targeting. He also felt that his research was urgent: new cases of MERS, spread by camels, were even then popping up in the Middle East. Eventually the NIH agreed, waving him forward.

No one can find the animal that gave people covid-19

Here's your guide to the WHO-China search for the origins of the coronavirus.

During a hearing on May 11, 2021, Senator Rand Paul confronted Anthony Fauci over funding of bat-virus research by the National Institutes of Health.

GREG NASH/POOL VIA AP

His 2015 paper, "A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence," was a tour de force, utilizing bleeding-edge genetic technology to alert the civilized world to a looming danger on its periphery. It also revived concerns about gain-of-function experiments, which Baric had known it would. In the paper, he spelled out the extra precautions he'd taken and held up the research as a test case. "The potential to prepare for and mitigate future outbreaks must be weighed against the risk of creating more dangerous pathogens," he wrote. "Scientific review panels may deem similar studies building chimeric viruses based on circulating strains too risky to pursue."

The NIH decided the risk was worth it. In a potentially fateful decision, it funded work similar to Baric's at the Wuhan Institute of Virology, which soon used its own reverse-genetics technology to make numerous coronavirus chimeras.

Unnoticed by most, however, was a key difference that significantly shifted the risk calculation. The Chinese work was carried out at biosafety level 2 (BSL-2), a much lower tier than Baric's BSL-3+.

What caused the covid-19 pandemic remains uncertain, and Shi says her lab never encountered the SARS-CoV-2 virus before the Wuhan outbreak. But now that US officials have said the possibility of a lab accident needs to be investigated, the spotlight has fallen on American funding of the Wuhan lab's less safe research. Todaya chorus of scientists, including Baric, are coming forward to say this was a misstep. Even if there is no link to covid-19, allowing work on potentially dangerous bat viruses at BSL-2 is "an actual scandal," says Michael Lin, a bioengineer at Stanford University.

The simmering concern that the US funded risky research in China burst into the national discussion on May 11, when Senator Rand Paul accused Anthony Fauci, the longtime director of the NIH's National Institute of Allergy and Infectious Diseases, of funding "supervirus" research in the US and "making a huge mistake" by trading the know-how to China. Paul repeatedly confronted Fauci and demanded to know if he had funded gain-of-function research in that country. Fauci denied the accusation, stating categorically: "The NIH has not ever, and does not now, fund gain-of-function research in the Wuhan Institute of Virology."

During a hearing on May 11, 2021, Senator Rand Paul confronted Anthony Fauci over funding of bat-virus research by the National Institutes of Health.

GREG NASH/POOL VIA AP

The denial rests on the NIH's specific definition of what was covered by the moratorium: work that would have deliberately enhanced SARS-like viruses, MERS, or flu by—for example—making them easier to spread through the air. The Chinese research did not have the specific goal of making the viruses more deadly, and rather than SARS itself, it used SARS's close cousins, whose real-world risk to humans was unknown—in fact, determining the risk was the point of the research. Just as when you trade in part of a poker hand for fresh cards, there was no way of knowing whether the final chimeras would be stronger or weaker.

The NIH has still not fully explained its decision-making and did not reply to questions. Citing a pending investigation, it has declined to release copies of the grant that sent the Wuhan institute about $600,000 between 2014 and 2019. It has also revealed little about its new system for assessing gain-of-function risks, which is carried out by an anonymous review panel whose deliberations are not made public. Until there's more sunlight, the agency will be fighting speculation, from Paul and others, that what occurred is a scenario Fauci himself had outlined in a 2012 commentary discussing research on pandemic germs.

"The only impact of this work is the creation, in a lab, of a new, non-natural risk."

Richard Ebright, Rutgers University

"Consider this hypothetical scenario," Fauci wrote. "An important gain-of-function experiment involving a virus with serious pandemic potential is performed in a well-regulated, world-class laboratory by experienced investigators, but the information from the experiment is then used by another scientist who does not have the same training and facilities and is not subject to the same regulations. In an unlikely but conceivable turn of events, what if that scientist becomes infected with the virus, which leads to an outbreak and ultimately triggers a pandemic?"

A wake-up call

Paul's grilling of Fauci brought new scrutiny to the relationship between Ralph Baric's lab at UNC and Zhengli Shi's at WIV, with some narratives painting Baric as the Sith master of SARS and Shi as his ascendant apprentice. They did share resources—for example, Baric sent the transgenic mice with human lung receptors to Wuhan. But after their initial collaboration, the two centers were more like competitors. They were in a race to identify dangerous coronaviruses, assess the potential threat, and develop countermeasures like vaccines.

Related Story

For Baric, that research started in the late 1990s. Coronaviruses were then considered low risk, but Baric's

studies on the genetics that allowed viruses to enter human cells convinced him that some might be just a few mutations away from jumping the species barrier.

That hunch was confirmed in 2002–'03, when SARS broke out in southern China, infecting 8,000 people. As bad as that was, Baric says, we dodged a bullet with SARS. The disease didn't spread from one person to another until about a day after severe symptoms began to appear, making it easier to corral through quarantines and contact tracing. Only 774 people died in that outbreak, but if it had been transmitted as easily as SARS-CoV-2, "we would have had a pandemic with a 10% mortality rate," Baric says. "That's how close humanity came."

Top researchers are calling for a real investigation into the origin of covid-19

A group of prominent biologists say there needs to be a "safe space" for asking whether the coronavirus came out of a lab.

As tempting as it was to write off SARS as a one-time event, in 2012 MERS emerged and began infecting people in the Middle East. "For me personally, that was a wake-up call that the animal reservoirs must have many, many more strains that are poised for cross-species movement," says Baric.

By then, examples of such dangers were already being discovered by Shi's team, which had spent years sampling bats in southern China to locate the origin of SARS. The project was part of a global viral surveillance effort spearheaded by the US nonprofit EcoHealth Alliance. The nonprofit—which has an annual income of over $16 million, more than 90% from government grants—has its office in New York but partners with local research groups in other countries to do field and lab work. The WIV was its crown jewel, and Peter Daszak, president of EcoHealth Alliance, has been a coauthor with Shi on most of her key papers.

By taking thousands of samples from guano, fecal swabs, and bat tissue, and searching those samples for genetic sequences similar to SARS, Shi's team began to discover many closely related viruses. In a cave in Yunnan Province in 2011 or 2012, they discovered the two closest, which they named WIV1 and SHC014.

Shi managed to culture WIV1 in her lab from a fecal sample and show that it could directly infect human cells, proving that SARS-like viruses ready to leap straight from bats to humans already lurked in the natural world. This showed, Daszak and Shi argued, that bat coronaviruses were a "substantial global threat." Scientists, they said, needed to find them, and study them, before they found us.

Many of the other viruses couldn't be grown, but Baric's system provided a way to rapidly test their spikes by engineering them into similar viruses. When the chimera he made using SHC014 proved able to infect human cells in a dish, Daszak told the press that these revelations should "move this virus from a candidate emerging pathogen to a clear and present danger."

To others, it was the perfect example of the unnecessary dangers of gain-of-function science. "The only impact of this work is the creation, in a lab, of a new, non-natural risk," the Rutgers microbiologist Richard Ebright, a longtime critic of such research, told Nature.

To Baric, the situation was more nuanced. Although his creation might be more dangerous than the original mouse-adapted virus he'd used as a backbone, it was still wimpy compared with SARS—certainly not the supervirus Senator Paul would later suggest.

In the end, the NIH clampdown never had teeth. It included a clause granting exceptions "if head of funding agency determines research is urgently necessary to protect public health or national security." Not only were Baric's studies allowed to move forward, but so were all studies that applied for exemptions. The funding restrictions were lifted in 2017 and replaced with a more lenient system.

Tyvek suits and respirators

If the NIH was looking for a scientist to make regulators comfortable with gain-of-function research, Baric was the obvious choice. For years he'd insisted on extra safety steps, and he took pains to point these out in his 2015 paper, as if modeling the way forward.

The CDC recognizes four levels of biosafety and recommends which pathogens should be studied at which level. Biosafety level 1 is for nonhazardous organisms and requires virtually no precautions: wear a lab coat and gloves as needed. BSL-2 is for moderately hazardous pathogens that are already endemic in the area, and relatively mild interventions are indicated: close the door, wear eye protection, dispose of waste materials in an autoclave. BSL-3 is where things get serious. It's for pathogens that can cause serious disease through respiratory transmission, such as influenza and SARS, and

the associated protocols include multiple barriers to escape. Labs are walled off by two sets of self-closing, locking doors; air is filtered; personnel use full PPE and N95 masks and are under medical surveillance. BSL-4 is for the baddest of the baddies, such as Ebola and Marburg: full moon suits and dedicated air systems are added to the arsenal.

“There are no enforceable standards of what you should and shouldn’t do. It’s up to the individual countries, institutions, and scientists.”

Filippa Lentzos, King’s College London

In Baric’s lab, the chimeras were studied at BSL-3, enhanced with additional steps like Tyvek suits, double gloves, and powered-air respirators for all workers. Local first-responder teams participated in regular drills to increase their familiarity with the lab. All workers were monitored for infections, and local hospitals had procedures in place to handle incoming scientists. It was probably one of the safest BSL-3 facilities in the world. That still wasn’t enough to prevent a handful of errors over the years: some scientists were even bitten by virus-carrying mice. But no infections resulted.

Brand-new pathogens

In 2014, the NIH awarded a five-year, $3.75 million grant to EcoHealth Alliance to study the risk that more bat-borne coronaviruses would emerge in China, using the same kind of techniques Baric had pioneered. Some of that work was to be subcontracted to the Wuhan Institute of Virology.

Related Story

They called it a conspiracy theory. But Alina Chan tweeted life into the idea that the virus came from a lab.

The whistleblowing scientist who advanced the lab-leak theory plans to change her name and disappear, but only after a book deal.

Two years later, Daszak and Shi published a paper reporting how the Chinese lab had engineered different versions of WIV1 and tested their infectiousness in human cells. The paper announced that the WIV had developed its own reverse-genetics system, following the Americans’ lead. It also included a troubling detail: the work, which was funded in part by the NIH grant, had been done in a BSL-2 lab. That meant the same viruses that Daszak was holding up as a clear and present danger to the world were being studied under conditions that, according to Richard Ebright, matched “the biosafety level of a US dentist’s office.”

Ebright believes one factor at play was the cost and inconvenience of working in high-containment conditions. The Chinese lab’s decision to work at BSL-2, he says, would have “effectively increas[ed] rates of progress, all else being equal, by a factor of 10 to 20”—a huge edge.

Work at the WIV was indeed progressing quickly. In 2017, Daszak and Shi followed with another study, also at BSL-2, that one-upped Baric’s work in North Carolina. The WIV had continued to unearth dozens of new SARS-like coronaviruses in bat caves, and it reported making chimeras with eight of them by fusing the spikes of the new viruses to the chassis of WIV1. Two of them replicated well in human cells. They were, for all intents and purposes, brand-new pathogens.

The revelation that the WIV was working with SARS-like viruses in subpar safety conditions has led some people to reassess the chance that SARS-CoV-2 could have emerged from some type of laboratory incident. “That’s screwed up,” the Columbia University virologist Ian Lipkin, who coauthored the seminal paper arguing that covid must have had a natural origin, told the journalist Donald McNeil Jr. “It shouldn’t have happened. People should not be looking at bat viruses in BSL-2 labs. My view has changed.”

But the WIV was not breaking any rules by working at BSL-2, says Filippa Lentzos, a biosecurity expert at King’s College London “There are no enforceable standards of what you should and shouldn’t do. It’s up to the individual countries, institutions, and scientists.” And in China, she says, the vertiginous rise of high-tech biological research has not been accompanied by an equivalent increase in oversight.

In an email, Zhengli Shi said she followed Chinese rules that are similar to those in the US. Safety requirements are based on what virus you are

studying. Since bat viruses like WIV1 haven't been confirmed to cause disease in human beings, her biosafety committee recommended BSL-2 for engineering them and testing them and BSL-3 for any animal experiments.

In response to questions about the decision to do the research in BSL-2 conditions, Peter Daszak forwarded a statement from EcoHealth Alliance stating that the organization "must follow the local laws of the countries in which we work" and that the NIH had determined the research was "not gain-of-function."

Questioning China

There is no law against using tighter lab security, however, and according to Baric, these viruses deserve it. "I would never argue that WIV1 or SHC014 should be studied at BSL-2, because they can grow in primary human cells," he says. "There's some risk associated with those viruses. We have no idea whether they could actually cause severe disease in a human, but you want to err on the side of caution ... If you study a hundred different bat viruses, your luck may run out."

Since the pandemic began, Baric has not said much about the possible origins of the virus or about his Chinese counterparts. On several occasions, however, he has quietly pointed to safety concerns at the WIV. In May 2020, when few scientists were willing to consider a lab leak in public, he published a paper acknowledging that "speculation about accidental laboratory escape will likely persist, given the large collections of bat virome samples stored in labs in the Wuhan Institute of Virology, the facility's proximity to the early outbreak, and the operating procedures at the facility." He flagged Daszak and Shi's BSL-2 paper, in case anyone didn't understand what he was saying.

Ralph Baric of the University of North Carolina specializes in the genetic engineering of coronaviruses as part of vaccine and drug research.

COURTESY PHOTO

The National Institutes of Health has also revisited its ties to the Wuhan lab. In April of 2020, the NIH terminated its grant to EcoHealth Alliance for bat virus research. In a follow-up letter to Daszak on July 8, it offered to reinstate the grant, but only if EcoHealth Alliance could allay its concerns, noting reports that the WIV "has been conducting research at its facilities in China that pose serious bio-safety concerns" for other countries. It added, "We have concerns that WIV has not satisfied safety requirements under the award, and that EcoHealth Alliance has not satisfied its obligations to monitor the activities of its subrecipient."

The genetic code of SARS-CoV-2 does not resemble that of any virus the WIV was known to be culturing in its lab, such as WIV1, and Baric says he still believes a natural spillover is the most likely cause. But he also knows the intricate risks of the work well enough to see a possible path to trouble. That is why, in May of this year, he joined 17 other scientists in a letter in the journal Science calling for a thorough investigation of his onetime collaborator's lab and its practices. He wants to know what barriers were in place to keep a pathogen from slipping out into Wuhan's population of 13 million, and possibly to the world.

"Let's face it: there are going to be unknown viruses in guano, or oral swabs, which are oftentimes pooled. And if you're attempting to culture a virus, you're going to have novel strains being dropped onto culture cells," Baric says. "Some will grow. You could get recombinants that are unique. And if that was being done at BSL-2, then there are questions you want to ask."

by Rowan Jacobsen

Exhibit

32

Implications of SARS Epidemic for China's Public Health Infrastructure and Political System

Testimony before the Congressional-Executive Commission on China
Roundtable on SARS
May 12, 2003

Yanzhong Huang
Assistant Professor of Political Science
Grand Valley State University*

The Return of the God of Plagues

Since November 2002, a form of atypical pneumonia called SARS (Severe Acute Respiratory Syndrome) has spread rapidly from China to Southeast Asia, Europe, and North America, prompting World Health Organization (WHO) to declare the ailment "a worldwide health threat." According to the organization, as of May 10, 2003, a cumulative total of 7,296 cases and 526 deaths have been reported from 33 countries or regions. The country that is particularly hit by the disease is China, where the outbreak of SARS has infected more than 4,800 people and killed at least 235 nationwide (excluding Hong Kong and Macao). The worst-hit city is China's capital Beijing, which has more than 2,200 cases - nearly half China's total - and 116 deaths. History is full of ironies: the epidemic caught China completely off guard forty-five years after Mao Zedong bade "Farewell to the God of Plagues."

The SARS epidemic is not simply a public health problem. Indeed, it has caused the most severe social-political crisis to the Chinese leadership since the 1989 Tiananmen crackdown. Outbreak of the disease is fueling fears among some economists that China's economy might be headed for a serious downturn. It already seems likely to wipe out economic growth in the second quarter and possibly reduce the growth rate for the entire year to about six percent, well below the level the government says it required to absorb millions of new workers who need jobs. The disease has also spawned anxiety, panic and rumour-mongering, which has already triggered a series of protests and riots in China.[1] Meanwhile, the crisis has underscored the tensions and conflicts among the top leadership, and undermined the government's efforts to create a milder new image in the international arena. As Premier Wen Jiabao pointed out in a recent cabinet meeting on the epidemic, at stake were "the health and security of the people, overall state of reform, development, and stability, and China's national interest and international image." How to manage the crisis has become the litmus test of the political will and ability of the fourth generation of Chinese leadership.

Given the political aspect of the crisis, this testimony will consider not only problems in China's public health infrastructure but also dynamics of its political system. It proceeds in three sections. The first section focuses on the making of the crisis, and discusses how problems in the health and political systems allowed SARS to transform from a sporadic nuisance to an epidemic that now affects hundreds of millions of people across the country. The next section considers the government crusade against SARS, and examines how the state capacity in controlling the disease is complicated and compromised by the health infrastructure and political system. The last section concludes with some policy recommendations for the Commission to consider.

The Making of A Crisis (November 2002-April 2003)

Information Blackout in Guangdong

With hindsight, China's health system seemed to respond relatively well to the emergence of the illness. The earliest case of SARS is thought to occur in Foshan, a city southwest of Guangzhou in Guangdong province, in mid-November 2002. It was later also found in Heyuan and Zhongshan in Guangdong. This "strange disease" alerted Chinese health personnel as early as mid-December. On January 2, a team of health experts were sent to Heyuan and diagnosed the disease as an infection caused by certain virus.[2] A Chinese physician, who was in charge of treating a patient from Heyuan in a hospital of Guangzhou, quickly reported the disease to local anti-epidemic station.[3] We have reason to believe that the local anti-epidemic station alerted the provincial health bureau about the disease, and the bureau in turn reported to the provincial government and the Ministry of Health (MoH) shortly afterwards, since the first team of experts sent by the Ministry arrived at Guangzhou on January 20 and the new provincial government (who took over on January 20) ordered an investigation of the disease almost at the same time.[4] A combined team of health experts from the Ministry and the province was dispatched to Zhongshan and completed an investigation report on the unknown disease. On January 27, the report was sent to the provincial health bureau and, presumably, Ministry of Health in Beijing. The report was marked "top secret," which means that only top provincial health officials could open it.

Further government reaction to the emerging disease, however, was delayed by the problems of information flow within the Chinese hierarchy. For three days, there were no authorized provincial health officials available to open the document. After the document was finally read, the provincial bureau distributed a bulletin to hospitals across the province. Yet few health workers were alerted by the bulletin, because most were on vocation for the Chinese New Year.[5] Meanwhile, the public was kept uninformed about the disease. According to the 1996 Implementing Regulations on the State Secrets Law (1988), any such diseases should be classified as a state secret before they are "announced by the Ministry of Health or organs authorized by the Ministry." In other words, until such time the Ministry chose to make public about the disease, any physician or journalist who reported on the disease would risk being persecuted for leaking state secrets.[6]

In fact, until February 11, not only news blackout continued, but the government failed to take any further actions on the looming catastrophe. Evidence indicated that the provincial government in deciding whether to publicize the event considered more about local economic development than about people's life and health. The Law on Prevention and Treatment of Infectious Diseases enacted in September 1989 contains some major loopholes. First, provincial governments only after being authorized by MoH are obliged to publicize epidemics in a timely and accurate manner (Article 23). Second, atypical pneumonia was not listed in the law as an infectious disease under surveillance, thus local government officials legally were not accountable for the disease. The law allows addition of new items to the list, but it does not specify the procedures through which new diseases can be added. All this provided disincentives for the government to effectively respond to the crisis.

To be sure, the media blackout and the government's slow response are not only the sole factors leading to the crisis. Scientists until today are still not entirely clear about the pathogen, spread pattern and mortality rate of SARS.[7] Due to the lack of knowledge about the disease, the top-secret document submitted to the provincial health bureau did not even mention that the disease was highly contagious, neither did it call for rigorous preventive measures, which may explain why by the end of February, nearly half of Guangzhou's 900 cases were health care workers.[8] Indeed, even rich countries, like Canada, were having difficulty controlling SARS. In this sense, SARS is a natural disaster, not a man-made one.

Yet there is no doubt that government inaction resulted in the crisis. To begin with, the security designation of the document means that health authorities of the neighboring Hong Kong SAR was not informed about the disease and, consequently, denied the knowledge they needed to prepare for

outbreaks.[9] Very soon, the illness developed into an epidemic in Hong Kong, which has proved to be a major transit route for the disease. Moreover, the failure to inform the public heightened anxieties, fear, and widespread speculation. On February 8, reports about a "deadly flu" began to be sent via short messages on mobile phones in Guangzhou. In the evening, words like bird flu and anthrax started to appear on some local Internet sites.[10] On February 10, a circular appeared in the local media acknowledged the presence of the disease and listed some preventive measures, including improving ventilation, using vinegar fumes to disinfect the air, and washing hands frequently. Responding to the advice, residents in Guangzhou and other cities cleared pharmacy shelves of antibiotics and flu medication. In some cities, even the vinegar was sold out. The panic spread quickly in Guangdong, and had it felt even in other provinces.

On February 11, Guangdong health officials finally broke the silence by holding press conferences about the disease. The provincial health officials reported a total of 305 atypical pneumonia cases in the province. The officials also admitted that there were no effective drugs to treat the disease, and the outbreak was only tentatively contained.[11] From then on until February 24, the disease was allowed to report extensively. Yet in the meantime, the government played down the risk of the illness. Guangzhou city government on February 11 went as far as to announce the illness was "comprehensively" under effective control.[12] As a result, while the panic was temporally allayed, the public also lost vigilance about the disease. During the run-up to the National People's Congress, the government halted most reporting. The news blackout would remain until April 2.

Beyond Guangdong: Ministry of Health and Beijing

Under the Law on Prevention and Treatment of Infectious Diseases, MoH is obliged to accurately report and publicize epidemics in time. The Ministry learned about SARS in January and informed WHO and provincial health bureaus about the outbreak in Guangdong around February 7. Yet no further action was taken. It is safe to assume that Zhang Wenkang, the health minister, brought the disease to the attention of Wang Zhongyu (Secretary General of the State Council) and Li Lanqing (the vice premier in charge of public health and education). We do not know what happened during this period of time; it is very likely that the leaders were so preoccupied preparing for the National People's Congress in March that no explicit directive was issued from the top until April 2.

As a result of the inaction from the central government and the continuous information blackout, the epidemic in Guangdong quickly spread to other parts of China. Since March 1, the epidemic has raged in Beijing. Yet for fear of disturbance during the NPC meeting, city authorities kept information about its scope not only from the public but also from the Party Center. MoH was reportedly aware of what was happening in the capital. The fragmentation of bureaucratic power, however, delayed any concerted efforts to address the problem. As one senior health official admitted, before anything could be done, the ministry had to negotiate with other ministries and government departments.[13] On the one hand, Beijing municipal government apparently believed that it could handle the situation well by itself and thus refused involvement of MoH. On the other hand, the Ministry did not have control of all health institutions. Of Beijing's 175 hospitals, 16 are under the control of the army, which maintains a relatively independent health system. Having admitted a large number of SARS patients, military hospitals in Beijing until mid-April refused to hand in SARS statistics to the Ministry. According to Dr. Jiang Yanyong, medical staff in Beijing's military hospitals were briefed about the dangers of SARS in early March, but told not to publicize what they had learned lest it interfere with the NPC meeting.[14] This might in part explain why on April 3, the health minister announced that Beijing had seen only 12 cases of SARS, despite the fact that in the city's No. 309 PLA hospital alone there were 60 SARS patients. The bureaucratic fragmentation also created communication problems between China and World Health Organization. WHO experts were invited by the Ministry to China but were not allowed to have access to Guangdong

until April 2, eight days after their arrival. They were not allowed to inspect military hospitals in Bejing until April 9. By that time, the disease had already engulfed China and spread to the world.

What is to blame?

The crisis revealed two major problems inherent in China's political system: cover-up and inaction. Existing political institutions have not only obstructed the information flow within the system but also distorted the information itself, making misinformation endemic in China's bureaucracy. Because government officials in China are all politically appointed rather than elected by the general populace at each level of administration, they are held accountable only to their superiors, not the general public. This upward accountability generates perverse incentives for government officials in policy process. For fear that any mishap reported in their jurisdiction may be used as an excuse to pass them over for promotion, government officials at all levels tend to distort the information they pass up to their political masters in order to place themselves in a good light. While this is not something unique to China, the problem is alleviated in democracies through "decentralized oversight," which enables citizen interest groups to check up on administrative actions. Since China still refuses to enfranchise the general public in overseeing the activities of government agencies, the upper-level governments are easier to be fooled by their subordinates. This exacerbates the information asymmetry problems inherent in a hierarchical structure and weakens effective governance of the central state.

Nevertheless, a functionalist argument can be made to explain the rampant underreporting and misreporting in China's officialdom. In view of the dying communist ideology and the official resistance to democracy, the legitimacy of the current regime in China is rooted in its constant ability to promote social-economic progress. As a result of this performance-based legitimacy, "government officials routinely inflate data that reflect well on the regime's performance, such as growth rates, while under reporting or suppressing bad news such as crime rates, social unrest and plagues."[15] In this sense, manipulation of data serves to shore up the regime's legitimacy.

In explaining the government's slow response to tackling the original outbreak, we should keep in mind that the health system is embedded in an authoritarian power structure in which policies are expected to come from the political leadership. In the absence of a robust civil society, China's policy making does not feature a salient "bottom-up" process to move a "systemic" agenda in the public to a "formal" or governmental agenda as found in many liberal democracies. To be sure, the process is not entirely exclusionary, for the party's "mass line" would require leading cadres at various levels to obtain information from the people and integrate it with government policy during the policy formation stage. Yet this upward flow of information is turned on or off like a faucet by the state from above, not by the strivings of people from below.[16] Under this top-down political structure, each level takes its cue from the one above. If the leadership is not dynamic, no action comes from the party-state apparatus. The same structure also encourages lower-level governments to shift their policy overload to the upper levels in order to avoid taking responsibilities. As a result, a large number of agenda items are competing for the upper level government's attention. The bias toward economic development in the reform era nevertheless marginalized the public health issues in the top leaders' agenda. As a matter of fact, prior to the SARS outbreak, public health had become the least of the concerns of Chinese leaders. Compared to an economic issue a public health problem often needs an attention-focusing event (e.g., a large-scale outbreak of a contagious disease) to be finally recognized, defined, and formally addressed. Not surprisingly, SARS did not raise the eyebrows of top decision makers until it had already developed into a nationwide epidemic.

Another problem that bogged down government response is bureaucratic fragmentation. Because Chinese decision-making emphasizes consensus, the bureaucratic proliferation and elaboration in the post-Mao era requires more time and effort for coordination. With the involvement of multiple actors in multiple

sectors, the policy outcome is generally the result of the conflicts and coordination of multiple sub-goals. Since units (and officials) of the same bureaucratic rank cannot issue binding orders to each other, it is relatively easy for one actor to frustrate the adoption or successful implementation of important policies. This fragmentation of authority is also worsened by the relationship between functional bureaucratic agency (*tiao*) and the territorial governments (*kuai*). In public health domain, territorial governments like Beijing and Guangdong maintain primary leadership over the provincial health bureau, with the former determining the size, personnel, and funding of the latter. This constitutes a major problem for the Ministry of Health, which is bureaucratically weak, not to mention that its minister is just an ordinary member of CCP Central Committee and not represented in the powerful Politburo. A major policy initiative from the Ministry of Health, even issued in the form of a central document, is mainly a guidance document (*zhidao xin wenjian*) that has less binding power than one that is issued by territorial governments. Whether they will be honored hinges on the "acquiescence" (*liangjie*) of the territorial governments. This helps explain the continuous lack of effective response in Beijing city authorities until April 17 (when the anti-SARS joint team was established).

China's Crusade against SARS (April 2003 – present)

Reverse Course

Thanks to strong international pressure, the government finally woke up and began to tackle the crisis seriously. On April 2, the State Council held its first meeting to discuss the SARS problem. Within one month, the State Council held three meetings on SARS. An order from the MoH in mid-April formally listed SARS as a disease to be monitored under the Law of Prevention and Treatment of Infectious Diseases and made it clear that every provincial unit should report the number of SARS on a given day by 12 noon on the following date. The party and government leaders around the country is now held accountable for the overall SARS situation in their jurisdictions. On April 17, an urgent meeting held by the Standing Committee of the Politburo explicitly warned against the covering up of SARS cases and demanded the accurate, timely and honest reporting of the disease. Meanwhile, the government also showed a new level of candor. Premier Wen Jiabao on April 13 said that although progress had been made, "the overall situation remains grave."[17] On April 20 the government inaugurated a nationwide campaign to begin truthful reporting about SARS.

The government also took steps to remove incompetent officials in fighting against SARS. Health minister Zhang Wenkang and Beijing mayor Meng Xuenong were discharged on April 20 to take responsibilities for their mismanagement of the crisis. While they were not the first ministerial level officials since 1949 who were sacked mid-crisis on a policy matter, the case did mark the first sign of political innovation from China's new leadership. According to an article in *Economist*, unfolding of the event (minister presides over policy bungle; bungle is exposed, to public outcry; minister resigns to take the rap) "almost looks like the way that politics works in a democratic, accountable country."[18] The State Council also sent out inspection teams to the provinces to scour government records for unreported cases and fire officials for lax prevention efforts. It was reported that since April, 120 government officials have lost their jobs.

The crisis also speeded up the process of institutionalizing China's emergency response system so that it can handle public health contingencies and improve interdepartmental coordination. On April 2, the government established a leading small group led by the health minister and an inter-ministerial roundtable led by a vice secretary general to address SARS prevention and treatment . This was replaced on April 23 by a task force known as the SARS Control and Prevention Headquarters of the State Council, to coordinate national efforts to combat the disease. Vice Premier Wu Yi was appointed as command-in-chief of the task force. On May 12, China issued Regulations on Public Health Emergencies (PHEs). According to the regulations, the State Council shall set up an emergency headquarters to deal with any

PHEs, which refer to serious epidemics, widespread unidentified diseases, mass food and industrial poisoning, and other serious public health threats. [19]

Meanwhile, the government increased its funding for public health. On April 23, a national fund of two billion *yuan* was created for SARS prevention and control. The fund will be used to finance the treatment of farmers and poor urban residents infected with SARS and to upgrade county-level hospitals and purchase SARS-related medical facilities in central and western China. The central government also committed 3.5 billion yuan for the completion of a three-tier (provincial, city, and county) disease control and prevention network by the end of this year. This includes 600 million for the initial phase of constructing China's Center for Disease Control and Prevention (CDC).[20] The government has also offered free treatment for poor SARS patients.

The government also showed more interest in international cooperation in fighting against SARS. In addition to its cooperation with WHO, China showed flexibility in cooperating with neighboring countries in combating SARS. At the special summit called by ASEAN and China in late April, Chinese premier Wen Jiabao pledged 10 million *yuan* to launch a special SARS fund and joined the regionwide confidence-building moves to take coordinated action against the disease.

Problems and Concerns

These measures are worth applauding, but are they going to work? The battle against the disease can be compromised by China's inadequate public health system. One of the major problems here is the lack of state funding. Already, the portion of total health spending financed by the government has fallen from 34 percent in 1978 to less than 20 percent now.[21] Cash-strapped local governments whose health-care system is underfinanced would be extremely hard pressed in the process of SARS prevention and treatment. It is reported that some hospitals have refused to accept patients who have affordability problems.[22] The offer of free treatment for poor SARS patients is little consolation to the large numbers with no health insurance, particularly the unemployed and the millions of ill-paid migrant workers, who are too poor to consider hospital treatment which getting sick. According to a 1998 national survey, about 25.6 percent of the rural patients cited "economic difficulties" as the main reason that they did not seek outpatient care.[23]

The lack of facilities and qualified medical staff to deal with the SARS outbreak also compromises government efforts to contain the disease. Among the 66,000 health care workers in Beijing, less than 3000, or 4.3 percent of them are familiar with respiratory diseases.[24] Similarly, hospitals in Guangdong are reported to face shortage in hospital beds and ambulances in treating SARS. This problem is actually worsened by the absence of referral system and the increasing competition between health institutions, which often leads to little coordination but large degrees of overlap. As SARS cases increases, some hospitals are facing the tough choice of losing money or not admitting further SARS patients. In Beijing, the government had to ask for help from the military.

Tremendous inequalities in health resource distribution posed another challenge to the Chinese leadership. To the extent that health infrastructure are strained in Beijing, the situation would be much worse in China's hinterland or rural areas. Compared with Beijing, Shanhai, and Jiangsu and Zhejiang provinces, which receives a full quarter of health-care spending, the seven provinces and autonomous regions in the far west only get 5 percent.[25] The rural-urban gap in health resource distribution is equally glaring. Representing only 20 percent of China's population, urban residents claim more than 50 percent of the country's hospital beds and health professionals. So far, a large-scale epidemic has not yet appeared in the countryside. The percentage of peasants who are infected, however, is high in Hebei, Inner Mongolia, and Shanxi, which points to the relatively high possibility of spread to the rural areas.[26]

Some other concerns also complicate the war on SARS. In terms of the mode of policy implementation, the Chinese system is in full mobilization mode now. All major cities are on 24-hour alert, apparently in response to emergency directions from the central leadership. So far, all indications point to decisive action for quarantine. By May 7, 18,000 people had been quarantined in Beijing. Meanwhile, the Maoist "Patriotic Hygiene Campaign" has been revitalized. In Guangdong, 80 million people were mobilized to clean houses and streets and remove hygienically dead corners.[27]By placing great political pressure on local cadres in policy implementation, mobilization is a convenient bureaucratic tool for overriding fiscal constraints and bureaucratic inertia whilst promoting grassroots cadres to behave in ways that reflect the priorities of their superiors. Direct involvement of the local political leadership increases program resources, helps ensure they are used for program purpose, and mobilizes resources from other systems, including free manpower transferred to program tasks. Yet in doing so a bias against routine administration was built into the implementation structure. In fact, the increasing pressure from higher authorities, as indicated by the system that holds government heads personally responsible for SARS spread under their jurisdiction, makes strong measures more appealing to local officials, who find it safer to be overzealous than to be seen as "soft." There are indications that local governments overkill in dealing with SARS. In some cities, those who were quarantined lost their jobs. Until recently, Shanghai was quarantining people from some regions hard hit by SARS (such as Beijing) for 10 days even if they had no symptoms.[28] While many people are cooperating with the government measures, there is clear evidence suggesting that some people were quarantined against their will.[29]

The heavy reliance on quarantine raises a question that should be of interest to the committee: will anti-SARS measures worsen human rights situations in China? This question of course is not unique to China: even countries like the U.S. are debating whether it is necessary to apply dictatorial approach to confront health risks more effectively. The Model Emergency Health Powers pushed by the Bush administration would permit state governors in a health crisis to impose quarantines, limit people's movements and ration medicine, and seize anything from dead bodies to private hospitals.[30] While China's Law on Prevention and Treatment of Infectious Disease does not explicate that quarantines apply to SARS epidemic, Articles 24 and 25 authorize local governments to take emergency measures that may compromise personal freedom. The problem is that unlike democracies, China in applying these measures excludes the input of civil associations. Without engaged civil society groups to act as a source of discipline and information for government agencies, the sate capability is often used not in the society's interest. Official reports suggested that innocent people were dubbed rumor spreaders and arrested simply because they relayed some SARS-related information to their friends or colleagues.[31] According to the Ministry of Public Security, since April public security departments have investigated 107 cases in which people used internet and cell phones to spread SARS-related "rumors."[32]Some Chinese legal scholars have already expressed concerns that the government in order to block information about the epidemic may turn to more human rights violations.[33]

The lack of engagement of civil society in policy process could deplete social capital so important for government anti-SARS efforts. As the government is increasingly perceived to be incapable of adequately providing the required health and other social services, it has alienated members of society, producing a heightened sense of marginalization and deprivation among affected populations. These alienated and marginalized people have even less incentive than they would ordinarily have to contribute to government-sponsored programs. The problem can be mitigated if workers and peasants are allowed to form independent organizations to fight for their interests. Unfortunately, China's closed political system offers few institutional channels for the disadvantaged groups to express their private grievances. The government failure to publicize the outbreak in a timely and accurate manner and the ensuing quick policy switch caused further credibility problems for the government. *Washington Post* reported a SARS patient who fled quarantine in Beijing because he did not believe that the government would treat his disease free of charge. This lack of trust toward the government contributed to the spread of rumors even after the government adopted a more open stance on SARS crisis. In late April, thousands of residents of

a rural town of Tianjin ransacked a building, believing it would be used to house ill patients with confirmed or suspected SARS, even though officials insisted that it would be used only as a medical observation facility to accommodate people who had close contacts with SARS patients and for travelers returning from SARS hot spots. Again, here the lack of active civilian participation exacerbated the trust problems. In initiating the project the government had done nothing to consult or inform the local people.[34] Opposition to official efforts to contain SARS was also found in a coastal Zhejiang province, where several thousand people took part in a violent protest against six people who were quarantined after returning from Beijing.[35]

Last but not least, policy difference and political conflicts within the top leadership can cause serious problems in polity implementation. The reliance on performance legitimacy put the government in a policy dilemma in coping with the crisis. If it fails to place the disease under control and allows it to run rampant, it could become the event that destroys the Party's assertions that it improves the lives of the people. But if the top priority is on health, economic issues will be moved down a notch, which may lead to more unemployment, more economic loss and more social and political instability. The disagreement over the relationship between the two was evidenced in the lack of consistence in official policy. On April 17, the CCP Politburo Standing Committee meeting focused on SARS. In a circular issued after the meeting, the Party Center made it clear that "despite the daunting task of reform and development, the top priority should be given to people's health and life security. We should correctly deal with the temporary loss in tourism and foreign trade caused by atypical pneumonia, have long-range perspective in thinking or planning, and do not concern too much about temporary loss."[36] Eleven days later, the Politburo meeting emphasized Jiang Zemin's "Three Represents" and, by calling for a balance between combating SARS and economic work, reaffirmed the central status of economic development.[37] This schizophrenic nature of central policy is going to cause at least two problems that will not help the state to boost its capacity in combating SARS. First, because the Party Center failed to signal its real current priorities loud and clear, local authorities may get confused and face a highly uncertain incentive structure of rewards and punishments. Given the central government's inability to perfectly differentiate between simple incompetence and willful disobedience, local policy enforcers may take advantage of the policy inconsistency to "shirk" or minimize their workload, making strict compliance highly unlikely. Second, the policy difference will aggravate China's faction-ridden politics, which in turn can reduce central leaders' policy autonomy so important for effectively fighting against SARS. A perceived crisis can precipitate state élites to fully mobilize the potential for autonomous action. Yet power at the apex in China inheres in individual idiosyncrasies rather than institutions. This lack of institutionalization at the top level, coupled with the pretensions of a centralized bureaucracy, sets the stage for a very constrained from of politics, limiting what passed as national politics to relations among the top elite. A general rule in Chinese elite politics is that policy conflicts will be interwoven with factionalism. Former President Jiang's allies in the Politburo Standing Committee seemed to be quite slow to respond to the anti-SARS campaign embarked on by Hu Jintao and Wen Jiabao on April 20. Wu Bangguo, Jia Qinglin, and Li Changchun did not show up on the front stage of SARS campaign until April 24. The absence of *esprit de corps* among key élites would certainly reduce state autonomy needed in handling the crisis. It is speculated that the fall of Meng Xuenong, a protégé of Hu, was to balance the removal of Zhang Wenkang, a Jiang follower. Given that a health minister, unlike a mayor of Beijing, is not a major power player, this seems to send a message that the former president is still very much in control. The making of big news Jiang's order on April 28 to mobilize military health personnel only suggests the lack of authority of Hu Jintao and Wen Jiabao over the military. Intraparty rivalry in handling the crisis reminded people political upheavals in 1989, when the leaders disagreed on how to handle the protests and Deng Xiaoping the paramount leader played the game between his top associates before finally siding with the conservatives by launching a military crackdown.

III. Policy Recommendations

The above analysis clearly points to the need for the Chinese government to beef up its capacity in combating SARS. Given that a public health crisis reduces state capacity when ever-increasing capacity is needed to tackle the challenges, purely endogenous solutions to build capacity are unlikely to be successful, and capacity will have to be imported from exogenous sources such as massive foreign aid.[38] In this sense, building state capability also means building more effective partnerships and institutions internationally. As I summarized somewhere else, international actors can play an important role in creating a more responsible and responsive government in China.[39] First, aid from international organizations opens an alternative source of financing health care, increasing the government's financial capacity in the health sector. Second, international aid can strengthen the bureaucratic capacity through technical assistance, policy counseling, and personnel training. Third, while international organizations and foreign governments provide additional health resources in policy implementation, the government increasingly has to subject its agenda-setting regime to the donors' organizational goals, which can make the government more responsive to its people. The recent agenda shift to a large extent was caused by the strong international pressures exerted by the international media, international organizations, and foreign governments. There is indication that Internet is increasingly used by the new leadership to solicit policy feedback, collect public opinions and mobilize political support. Starting February 11, Western news media were aggressively reporting on SARS and on government cover-up of the number of cases in China. It is very likely that Hu Jintao and Wen Jiaobao, both Internet users, made use of international information in making decisions on SARS. In other words, external pressures can be very influential because Chinese governmental leaders are aware of the weakness of the existing system in effectively responding to the crisis, and have incentives to seek political resources exogenous to the system.

From the perspective of international actors, helping China fighting SARS is also helping themselves. Against the background of a global economy, diseases originating in China can be spread and transported globally through trade, travel, and population movements. Moreover, an unsustainable economy or state collapse spawned by poor health will deal a serious blow to the global economy. As foreign companies shift manufacturing to China, the country is becoming a workshop to the world. A world economy that is so dependent on China as an industrial lifeline can become increasingly vulnerable to a major supply disruption caused by SARS epidemic. Perhaps equally important, if the SARS epidemic in China runs out of control and triggers a global health crisis, it will result in some unwanted social and political changes in other countries including the United States. As every immigrant or visit from China or Asia is viewed as a Typhoid Mary, minorities and immigration could become a sensitive domestic political issue. The recent incident in New Jersey, in which artists with Chinese background were denied access to a middle school, suggests that when SARS becomes part of a national lexicon, fear, rumor, suspicion, and misinformation can jeopardize racial problems in this country.[40]

Given the international implications of China's public health, it is in the U.S. interest to expand cooperation with China in areas of information exchange, research, personnel training, and improvement of public health facilities. But it can do more. It can modify its human rights policy so that it accords higher and clearer priority to health status in China. Meanwhile, it could send a clearer signal to the Chinese leadership that the United States supports reform-minded leaders in the forefront of fighting SARS. To the extent that regime change is something the U.S. would like to see happening in China, it is not in the U.S. interest to see Hu Jingtao and Wen Jiaobao purged and replaced by a less open and less humane government, even though that government may still have strong interest in maintaining a healthy U.S.-China relationship. The United States simply should not miss this unique opportunity to help create a healthier China.

* Beginning September 2003, the author will be an assistant professor of the John C. Whitehead School of Diplomacy and International Relations at Seton Hall University and inaugural director of Global Health Studies Center.

---

[1] Anthony Kuhn, "China's Fight Against SARS Spawns Backlash," *Los Angeles Times*, My 6, 2003.

[2] "Guangzhou is fighting an unknown virus," *Southern Weekly*, February 13, 2003.

[3] Renmin ribao, overseas edition, 22 April 2003.

[4] http://www.people.com.cn/GB/shehui/47/20030211/921420.html.

[5] John Pomfret, "China's slow reaction to fast-moving illness," *Washington Post*, 3 April 2003, p. A18.

[6] Li Zhidong, et al, *Zhonghua renmin gonghe guo baomifa quanshu* (Encyclopedia on the PRC State Secrets Law) (Changchun: Jilin renmin chubanshe, 1999), pp. 372-374. I thank Professor Richard Baum for bringing this to my attention.

[7] On February 18, the Chinese CDC identified chlamydia bacteria as the cause of the disease. At the end of the month, WHO experts believed the disease was an outbreak of bird flue. They did not identify it as a new infectious disease until early March.

[8] Pomfret, "China's slow reaction to fast-moving illness."

[9] *Ibid*.

[10] *South China Morning Post*, February 11, 2003.

[11] *Southern Weekly*, February 13, 2003.

[12] http://www.people.com.cn/GB/shehui/47/20030211/921422.html.

[13] John Pomfret, "China's Crisis Has a Political Edge," *Washington Post*, April 27, 2003.

[14] Susan Jakes, "Beijing's SARS Attack," *Time*, April 8, 2003.

[15] Minxin Pei, "A Country that does not take care of its people," *Financial Times*, April 7, 2003.

[16] Jean Oi, *State and Peasant in Contemporary China* (Berkeley: University of California Press, 1989), p. 228.

[17] *BusinessWeek*, April 28, 2003.

[18] "China's Chernobyl," *Economist*, April 26, 2003, p. 9

[19] Xinhua News, http://news.xinhuanet.com/newscenter/2003-05/12/content_866362.htm.

[20] *Renmin ribao* (People's daily), overseas edition, May 9, 2003.

[21] Yanzhong Huang, *Mortal Peril: Public Health in China and Its Security Implications*. CBACI Health and Security Series, Special Report 6, May 2003.

[22] *Washington Post*, April 14, 2003.

[23] Ministry of Health, *National Health Service Research*. Beijing, 1999.

[24] *Renmin ribao*, overseas edition, May 1, 2003.

[25] *BusinessWeek*, April 28, 2003.

[26] *Xinhua News*, May 10, 2003.

[27] *Renmin ribao*, April 9, 2003.

[28] Pomfret, "China Feels Side Effects from SARS," *Washington Post*, May 2, 2003.

[29] *Beijing Youth Daily*, May 2, 2003; http://www.people.com.cn/GB/shehui/45/20030510/988713.html

[30] Nicholas D. Kristof, "Lock 'Em Up," *New York Times*, May 2, 2003.

[31] http://www.people.com.cn/GB/shehui/47/20030426/980282.html;

[32] http://www.people.com.cn/GB/shehui/44/20030508/987610.html. May 8, 2003.

[33] http://www.duoweinews.com Accessed on May 10, 2003.

[34] Erik Eckholm, "Thousands Riot in Rural Chinese Town over SARS," *New York Times*, April 28, 2003.

[35] "China's fight against SARS spawns backlash," *Los Angeles Times*, May 6, 2003.

[36] http://www.people.com.cn/GB/shizheng/3586/20030422/977907.html, April 22, 2003

[37] *Renmin ribao*, April 29, 2003

[38] Andrew T. Price-Smith, "Pretoria's Shadow: The HIV/AIDS Pandemic and National Security in South Africa," Special Report No. 4, CBACI Health and Security Series, September 2002, p. 27.

[39] *Mortal Peril: Public Health in China and Its Security Implications.*

[40] "Fear, not SARS, rattles South Jersey School," *New York Times*, May 10, 2003.

Exhibit

33

www.nature.com/scientificreports

scientific reports

Check for updates

OPEN

In silico comparison of SARS-CoV-2 spike protein-ACE2 binding affinities across species and implications for virus origin

Sakshi Piplani[1,2], Puneet Kumar Singh[2], David A. Winkler[3,4,5]✉ & Nikolai Petrovsky[1,2]✉

The devastating impact of the COVID-19 pandemic caused by SARS–coronavirus 2 (SARS-CoV-2) has raised important questions about its origins and the mechanism of its transfer to humans. A further question was whether companion or commercial animals could act as SARS-CoV-2 vectors, with early data suggesting susceptibility is species specific. To better understand SARS-CoV-2 species susceptibility, we undertook an in silico structural homology modelling, protein–protein docking, and molecular dynamics simulation study of SARS-CoV-2 spike protein's ability to bind angiotensin converting enzyme 2 (ACE2) from relevant species. Spike protein exhibited the highest binding to human (h)ACE2 of all the species tested, forming the highest number of hydrogen bonds with hACE2. Interestingly, pangolin ACE2 showed the next highest binding affinity despite having a relatively low sequence homology, whereas the affinity of monkey ACE2 was much lower despite its high sequence similarity to hACE2. These differences highlight the power of a structural versus a sequence-based approach to cross-species analyses. ACE2 species in the upper half of the predicted affinity range (monkey, hamster, dog, ferret, cat) have been shown to be permissive to SARS-CoV-2 infection, supporting a correlation between binding affinity and infection susceptibility. These findings show that the earliest known SARS-CoV-2 isolates were surprisingly well adapted to bind strongly to human ACE2, helping explain its efficient human to human respiratory transmission. This study highlights how in silico structural modelling methods can be used to rapidly generate information on novel viruses to help predict their behaviour and aid in countermeasure development.

The devastating impact of COVID-19 infections caused by SARS–coronavirus 2 (SARS-CoV-2) has stimulated unprecedented international activity to discover effective coronavirus vaccines and drugs[1–4]. It has also raised important questions, including the mechanisms of zoonotic transfer of these viruses from animals to humans, whether companion animals or those used for commercial purposes can act as infection reservoirs, and why large variations in susceptibility are seen across animal species[5–7]. Understanding how such viruses move between species may help us prevent or minimize similar events in the future. Methods that elucidate the molecular basis for susceptibility differences may also help explain why different human populations exhibit different susceptibilities[8].

The spike protein (S protein) with its functional polybasic furin cleavage site at the S1–S2 boundary plays a key role in SARS-CoV-2 infectivity[9]. The S protein monomer consists of a fusion peptide, two heptad repeats, an intracellular domain, N-terminal domain, two subdomains and a transmembrane region[10]. Angiotensin converting enzyme 2 (ACE2) was identified as the main receptor for SARS-CoV-2 S protein, mimicking ACE2's role as the receptor for SARS virus. The binding of S protein to ACE2 is a critical initiating event for infection and human to human transmission (Fig. 1). ACE2 is relatively ubiquitously expressed in humans, being present in the lungs, arteries, heart, kidney, and intestines. ACE2 consists of an N-terminal peptidase M2 domain and a C-terminal collectrin renal amino acid transporter domain.

Non-human species vary markedly in their susceptibility to SARS-CoV-2[7,11,12]. As ACE2 sequences differ between species, this raises the possibility that differences in S protein's ability to bind ACE2 from different

[1]College of Medicine and Public Health, Flinders University, Bedford Park 5046, Australia. [2]Vaxine Pty Ltd, 11 Walkley Avenue, Warradale 5046, Australia. [3]Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, VIC 3086, Australia. [4]Monash Institute of Pharmaceutical Sciences, Monash University, Parkville 3052, Australia. [5]School of Pharmacy, University of Nottingham, Nottingham NG7 2RD, UK. ✉email: d.winkler@latrobe.edu.au; nikolai.petrovsky@flinders.edu.au

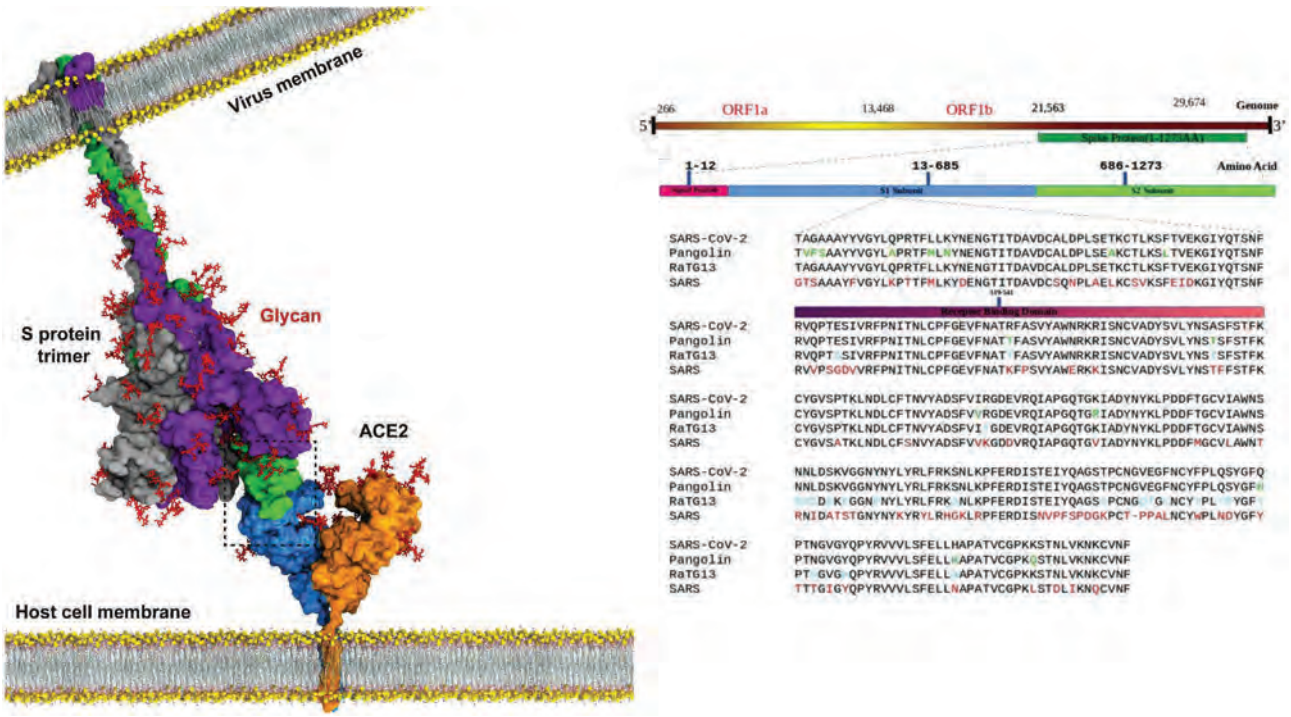


Figure 1. Upper panel: Picture of the complex formed by SARS-CoV-2 S protein and human ACE2 (CC-BY-NC-ND 4.0 International license; adapted from Taka et al.[15]). Lower panel: Sequence alignment of S1 subunit of four closely related spike proteins (sequence commences 60 residues before RBD). Coloured differences in spike protein sequences from pangolin CoV (green), bat RATG13 CoV (cyan) SARS CoV (red) when compared to sequence of SARS-CoV-2. The spike protein receptor binding domain (RBD) is denoted by the red bar in the sequence alignments.

species might underlie species susceptibility to infection. A phylogenetic tree showing ACE2 sequence relatedness across different animal species including bat, pangolin, and snake, all of which had been postulated as a source or intermediate host for SARS-Cov-2[11,13,14] is shown in Supplementary Fig. 1.

We and others[16–18] have postulated that structural variation in different species of ACE2 might determine S protein binding and thereby determine which species are permissive to SARS-CoV-2 infection[19]. For example, the low binding affinity of S protein for mouse ACE2 likely explains why mice are not susceptible to SARS-CoV2 infection. Direct measurement of the binding affinity of SARS-CoV-2 S protein to ACE2, e.g., using cell lines transfected with ACE2 proteins from different species, would provide valuable data but is time consuming, and purified or recombinant ACE2 proteins from many relevant animal species were not available at the time our modelling study was performed in early 2020. Hence we show here how an alternative approach using in silico structural modelling and docking algorithms from structure-based drug design was used to determine and compare the binding affinity of SARS-CoV-2 S protein to ACE2 of common and exotic animal species[20–22]. The species studied had either been implicated in transfer of SARS-CoV-2 to humans (e.g., bat, snake pangolin), been reported to be susceptible or resistant to SARS-CoV-2 infection (tiger, mouse, ferret, hamster, civet, monkey) or are important agricultural species or companion animals (cow, horse, cat, dog). The results provided novel insights into the species-specific nature of S protein-ACE2 interaction and predicted which species might be permissive for infection. They may also provide important clues into the origin of the pandemic since the mechanisms for SARS-CoV-2's appearance in the human population remains unknown despite more than a year passing since the start of the pandemic.

Results and discussion

SARS-Cov-2 S protein binds ACE2 of diverse species. The results of docking the receptor binding domains (RBDs) of SARS-CoV-2 S protein and ACE2 of various species using the HDOCK server, refined by MD simulations, are summarized in Tables 1 and 2. The calculated binding energies for the interactions are summarized in Table 2 with the MMPBSA binding energies listed for comparison. These are also presented graphically in Supplementary Fig. 3.

The energies calculated by Eq. 1 and those from the MMPBSA algorithm (see Materials and Methods) were strongly correlated, with $r^2 = 0.76$ (Supplementary Fig. 3). The Kendall tau rank correlation for the two methods of calculating binding energies was 0.94. Interestingly, the calculated energies correlated poorly with the degree of ACE2 sequence similarity ($r^2 = 0.27$), suggesting ACE2 structural, rather than sequence-based features have a dominant effect on S protein binding. Notably, docking and MD simulation allows binding affinities to be

Species	Accession number	Position																% binding residues shared with hACE2
		19	24	27	28	30	31	34	37	38	41	42	79	83	330	353	393	
Homo sapiens (human)	Q9BYF1	S	Q	T	F	D	K	H	E	D	Y	Q	L	Y	N	K	R	100
Macaca fascicularis (monkey)	A0A2K5X283	S	Q	T	F	D	K	H	E	D	Y	Q	L	Y	N	K	R	100
Panthera tigris (tiger)	XP_007090142.1	S	L	T	F	D	K	H	E	**E**	Y	Q	L	Y	K	K	R	94
Bos Taurus (cow)	NP_001019673.2	S	Q	T	F	**E**	K	H	E	D	Y	Q	M	Y	N	K	R	88
Mesocricetus auratus (hamster)	A0A1U7QTA1	S	Q	T	F	D	L	Q	E	D	Y	Q	L	Y	N	K	R	88
Felis catus (cat)	Q56H28	S	L	T	F	**E**	K	H	E	**E**	Y	Q	L	Y	N	K	R	81
Rhinolophus sinicus (bat)	U5WHY8	S	*E*	**M**	F	D	K	T	E	D	H	Q	L	Y	N	K	R	75
Paguma larvata (civet)	Q56NL1	S	L	T	F	**E**	K	Y	E	**Q**	Y	Q	L	Y	N	K	R	75
Equus ferus caballus (horse)	F6V9L3	S	L	T	F	D	K	S	E	**E**	H	Q	L	Y	N	K	R	75
Mustela putorius furo (ferret)	Q2WG88	D	L	T	F	**E**	K	T	E	**E**	Y	Q	-	Y	N	K	R	69
Canis luparis (dog)	J9P7Y2	–	L	T	F	**E**	K	Y	E	**E**	Y	Q	L	Y	N	K	R	69
Mus musculus (mouse)	Q8R0I0	S	**N**	T	F	**N**	N	Q	E	D	Y	Q	Y	F	N	K	R	63
Manis javanica (pangolin)	XP_017505752.1	–	*E*	T	F	**E**	K	S	E	**E**	Y	Q	**I**	Y	N	K	R	63
Ophiophagus Hannah (snake)	ETE61880.1	Q	V	K	F	**E**	Q	A	–	D	Y	**N**	N	**F**	N	L	R	38

Table 1. ACE2 RBD residues interacting with the S protein RBD from MD simulations of complexes. Residues interacting with the same S residue in different species ACE2, that differ from those in human ACE2, are in bold (conservative replacements), in italics (partially conservative replacements) or in underline (non-conservative replacements).

inferred. The 3D arrangement of amino acids determines the binding of two proteins with contribution from a broad range of surface features, including spatial and physicochemical (e.g., electrostatic and lipophilic) features. The present study integrates analysis of these features to determine the predicted binding affinity. Given the large numbers of models involved, with the exception of S protein complexes with human, monkey and pangolin ACE2, we did not analyse the individual molecular interactions contributing to the overall binding energy in each complex. The key interacting residues of ACE2 and spike proteins identified in our study, by inspecting the converged 3D ACE2-S protein complex structure from MD simulations (Table 1), were broadly consistent with other studies[32–34]. Interacting residues were identified using UCSF Chimera ← Find Contacts. Key S protein-interacting residues that were conserved across all species of ACE2 included PHE28, ASN330, LYS353 and ARG357. Other S protein-interacting residues in ACE2, namely TYR41, LYS353, ALA386 and ARG393, were conserved across all ACE2 species except bat, mouse, ferret and pangolin. S protein-interacting residues with King cobra ACE2 were the least conserved with human ACE2, consistent with the low sequence similarity of snake and human ACE2.

SARS-Cov-2 S protein optimally binds human ACE2. The ability to ascertain which species are permissive to infection could help identify potential intermediate hosts through which the SARS-CoV-2 virus crossed from a speculated bat source to humans. Although the SARS-CoV-2 S protein has only 72% sequence identity to the SARS receptor binding domain (RBD) region, like SARS-CoV-2, human, civet and bat SARS virus all use ACE2 for cellular entry[35–37]. The closest known relative to SARS-CoV-2 identified so far is the bat RaTG13 virus. However, RaTG13 has a different S protein to SARS-CoV-2 as it lacks a polybasic furin cleavage site and has major amino acid differences in its RBD that, at the sequence level, is most similar to the RBD of pangolin CoV (sequence data in Fig. 1). This has led to suggestions that pangolins might have served as the original host of SARS-CoV-2. Notably, our structure-based analysis of the species specificity of SARS-CoV-2 revealed some surprisingly results that differed from those from purely sequence based analyses. Conspicuously, we found that the binding of the SARS-CoV-2 S protein was higher for human ACE2 than any other species we tested, with the ACE2 binding energy order, from highest to lowest, being human > pangolin > dog > monkey > hamster > ferret > cat > tiger > bat > civet > horse > cow > snake > mouse. At its extremes, this ranking accords with experimental observations that humans are highly permissive to SARS-CoV-2 infection whereas mice on the other hand are not susceptible.

Potential significance of high S protein binding to human ACE2. During the early pandemic, S protein mutations were rare, especially in the RBD region interacting with ACE2. Notably, three mutation sites, V367F, G476S, and V483A were identified within the RBD of some SARS-CoV-2 isolates but only G476S was in the binding interface and its incidence and geographic spread was very small[38]. This minimal S protein RBD mutation during the early pandemic supports the view that the SARS-CoV-2 S protein was already optimally adapted for human ACE2 binding. This finding was surprising as a zoonotic virus typically exhibits the highest affinity initially for its original host species, with lower initial affinity to receptors of new host species until it adapts. As the virus adapts to its new host, mutations are acquired that increase the binding affinity for the new host receptor. Since our binding calculations were based on SARS-CoV-2 samples isolated in China from December 2019, at the very onset of the outbreak, the extremely high affinity of S protein for human ACE2 was unexpected. This high affinity was confirmed in July 2020 by Alexander et al.[39] who similarly found that

Species	ΔG_{eqn1} (kcal/mol)	ΔG_{MMPBSA} (kcal/mol)	SARS-Cov-2 infectivity
Homo sapiens (human)	−52.8	−57.6 ± 0.25	Permissive, high infectivity, severe disease in 5–10%,
Manis javanica (pangolin)	−52.0	−56.3 ± 0.4	Permissive[23,24]
Canis luparis (dog)	−50.8	−49.5	Permissive, low/mod infectivity, no overt disease[25,26]
Macaca fascicularis (monkey)	−50.4	−50.8	Permissive, high infectivity, lung disease[11]
Mesocricetus auratus (hamster)	−49.7	−50.0	Permissive, high infectivity, lung disease[27,28]
Mustela putorius furo (ferret)	−48.6	−49.2	Permissive, moderate infectivity, no overt disease[28–30]
Felis catus (cat)	−47.6	−48.9	Permissive, high infectivity, lung disease[26,29,31]
Panthera tigris (tiger)	−47.3	−42.5	Permissive, overt disease, RNA positive[26]
Rhinolophus sinicus (bat)	−46.9	−50.1 ± 1.0	Not permissive[11]
Paguma larvata (civet)	−45.1	−46.1	No reported infection
Equus ferus caballus (horse)	−44.1	−49.2	No naturally occurring infections[26]
Bos taurus (cow)	−43.6	−42.5	No naturally occurring infections[26]
Ophiophagus hannah (king cobra)	−39.5	−40.7 ± 1.2	No reported infection
Mus musculus (mouse)	−38.8	−39.4	Resistant to infection[28]

Table 2. Binding free energies of SARS-Cov-2 spike to ACE2 for different species and infection susceptibility reported by other studies.

the SARS-CoV-2 S protein RBD is optimal for binding to human ACE2 compared to other species. They also commented upon this as a remarkable finding that likely underlies the high transmissibility of SAR-Cov-2 virus among humans. Our results are also consistent with a study that found SARS-CoV-2 RBD bound with much higher affinity to hACE2 through additional hydrogen bonds and hydrophobic interactions when compared to binding of the SARS virus RBD to hACE2[34].

Possible role of pangolins as an intermediate host for SARS-CoV-2? Interestingly, as shown in Table 1, pangolin and human ACE2 are closest in spike binding energy despite being structurally different and only sharing 10 of 16 interacting residues at the SARS-CoV-2 RBD. The similarity in binding energy is noteworthy as pangolins have previously been imputed as a potential intermediate host to explain the spill-over of a putative bat coronavirus to humans. However, pangolin ACE2 was predicted to have significantly lower binding affinity to S protein than human ACE2 ($p = 0.0013$). The root-mean-square deviations (RMSD) between the simulated S protein-pangolin ACE2 and S protein-human ACE2 was just 1.211 Å (Fig. 2), indicating the close similarity of these complexes. Most of the pangolin ACE2 differences are conservative replacements of residues in human ACE2 viz., Q24E, D30E, D38E, and L79I, that are likely to make similar contributions to the binding interaction with SARS-CoV-2 spike. Of these, Q24E is only partially conservative because of the change in net charge. Q and E are similar based on the small physicochemical distances, chemical similarity, and experimental exchangeability (Grantham's distance based on composition, polarity and molecular volume; Epstein's coefficient of difference that is based on the differences in polarity and size between replaced pairs of amino acids; Miyata's distance based on volume and polarity; and Experimental Exchangeability devised by Yampolsky and Stoltzfus). E24 also interacts with an Asn residue in the spike RBD suggesting that hydrogen bonding ability rather than salt bridge formation is a more important contribution to the binding energy than electrostatic interactions characterizing salt bridges.

The key differences are the lack of S19 interaction and the replacement of H34 by S34 in pangolin ACE2. However, the MD structures show that the OH moiety in the sidechain of S34 in pangolin ACE2 lies in the same region, and can make similar interactions, as the NH moiety in the imidazole sidechain of H34 in human ACE2. Overall, given the high binding energy of S protein for pangolin ACE2, the possibility of pangolins being an intermediary vector for SARS-CoV-2 cannot be excluded (see also discussion below).

Exclusion of palm civets as a likely intermediate host for SARS-CoV-2. We were interested to know whether SARS-CoV-2 S protein could bind ACE2 from other species, notably the palm civet, the intermediate host for SARS virus[40]. The predicted binding affinity of SARS-CoV-2 S protein for palm civet ACE2 was low (ΔG_{MMPBSA} -46.1, > 10 kcal/mol lower than hACE2), making it extremely unlikely palm civets acted as an intermediate host for SARS-CoV-2.

Explanation for the large difference in S protein binding to human and monkey ACE2. A surprising observation, not identified by ACE2 sequence analysis, was that the binding affinity of S protein to monkey ACE2 (mACE2) was lower than to human ACE2 by 2.4—6.8 kcal kcal/mol (depending on the binding energy calculation method) despite all 16 S protein binding residues being shared between monkey and human ACE2. This suggested to us that binding energy differences must reside in structural differences rather than these specific residues.

The structure-based alignment of S protein RBD complexed with hACE2 and mACE2 after MD simulation exhibited an RMSD of 1.359 Å, highlighting differences between the 3-D structures of the two species despite their high sequence similarity (Fig. 3). Although almost all residues in the S protein binding site were common

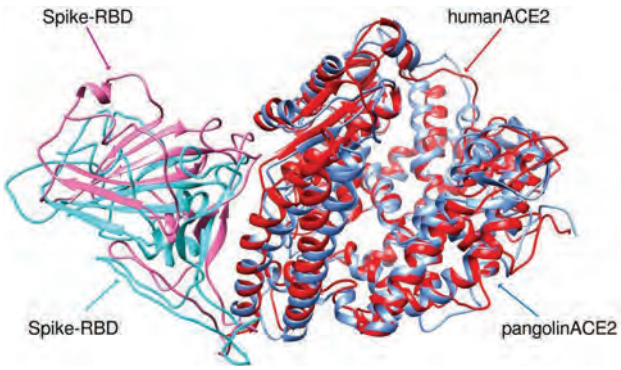


Figure 2. RMSD of overlay of S protein RBD (pink = with pangolin and turquoise = with human) complex with human ACE2 (red) or pangolin ACE2 (blue) after MD simulation showing different geometry of the two complexes.

to human and monkey ACE2 sequences (Table 1), structural rearrangements in the S protein and hACE2 binding complex resulted in more H-bonds forming between S protein and hACE2 than the equivalent complex with mACE2. As the binding of S protein to ACE2 is primarily governed by hydrogen bonds and electrostatic interactions, we monitored the total number of intermolecular H-bonds between the S protein and hACE2 and mACE2 proteins throughout the MD simulations. This revealed that the loop comprised of Thr478, Pro479, Cys480, Asn481, Gly482, Val483 and Glu484 plays an important role in the orientation and structural differences of human or monkey ACE2 bound to S protein. Although there is high conservation in the protein sequence of the interacting residues between human and monkey ACE2, MD simulations showed S protein formed more H-bonds with hACE2 (~ 640) than mACE2 (~ 620), thereby explaining the weaker binding affinity of S protein to mACE2, despite the high sequence similarity. Unlike hACE2, mACE2 specifically didn't form H-bonds with S protein residues Ser19, Phe28 and Lys353. Notably, a stable salt bridge was seen to form between Lys417 in S protein with Asp30 in hACE2 whereas this salt bridge was missing in the complex with mACE2. Lys353 in the S protein also formed an intermolecular salt bridge with Asp38 in hACE2 and was buried in a hydrophobic environment. Hence, although our model supports monkeys being permissive to SARS-CoV-2, it would predict that they should be less susceptible to severe clinical disease than humans. This marries well with experimental data which shows that monkeys are amongst species reported to be susceptible to SARS-CoV-2[5,29,41], with macaques, hamsters and ferrets being utilised as models of infection[27,30,42]. However, infected young cynomolgus macaques, while they expressed viral RNA in nasal swabs, did not develop overt clinical symptoms with aged animals exhibiting higher viral RNA loads, some weight loss, and moderate interstitial pneumonia and respiratory tract virus replication but ultimately spontaneously clearing the virus without treatment[42,43].

SARS-CoV-2 susceptibility for laboratory species. Our model showed hamster ACE2 had high S protein binding, similar to the binding affinity to monkey ACE2. This could be explained by hamster ACE2 sharing 14 out of 16 of the S protein binding residues seen in hACE2. Hence our model predicts hamsters should be permissive to SARS-CoV-2 infection (Table 2). In support of these findings, Syrian hamsters have been shown to exhibit clinical and histopathological responses to SARS-CoV-2 that closely mimic human respiratory tract infections, with high virus shedding and ability to transmit the virus to naïve contact animals[27]. The high susceptibility of hamsters to SARS-CoV-2 infection has made them one of the most utilised small animal models. Our model also predicted that ferret ACE2 has a similar binding strength to the S protein as hamster ACE2 (Table 2). Again, ferrets have been found permissive to SARS-CoV-2 infection, with high virus titre in the upper respiratory tract, virus shedding, acute bronchiolitis and active virus transmission to naïve ferrets through direct contact[29,30].

At the other extreme, based on the model predicting low S protein binding affinity for murine ACE2, mice should be resistant to infection. Supporting this finding, SARS-CoV-2 has been shown to have inefficient replication in mice[44]. Mice, however, become permissive for SARS or SARS-CoV-2 infection when made transgenic for human ACE2[45].

SARS-CoV-2 susceptibility of companion animals. Companion animals are in close contact with their human owners, creating a high risk of cross transmission if these animals were susceptible to SARS-CoV-2 infection. Feline ACE2 (cat and tiger) were shown by our model to have moderate to strong S protein binding affinity. Notably, both these species have now recognised as permissive for SARS-CoV-2 infection. Similarly, our data showed moderate to high affinity of S protein for dog ACE2, predicting dogs should be susceptible to infection. Again, Shen et al. found that that SARS-CoV-2 could be efficiently transmitted to both cats and dogs[46]. Shi et al. reported that ferrets and cats were more permissive to infection than dogs, pigs, chickens, and ducks[29]. Temmam et al. tested 9 cats and 12 dogs living in close contact with their owners with 2 testing positive for SARS-CoV-2 and 11 of 18 others showing clinical signs of COVID-19 but with no serum SARS-CoV-2 antibodies detectable[41]. Interestingly, Goumeniu et al. published an editorial querying the role of dogs in the Lombardy COVID-19 outbreak and recommended use of computational docking experiments, like our own, to

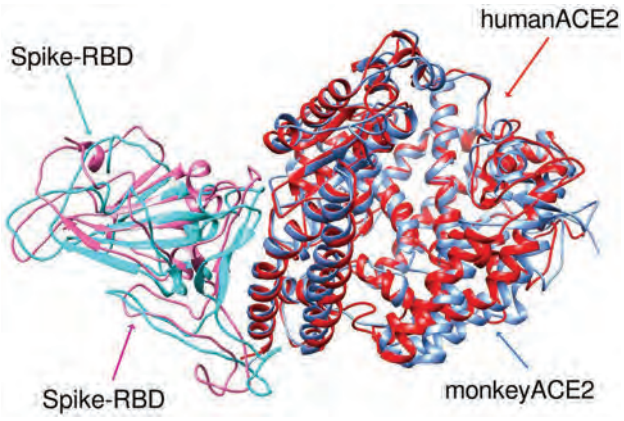


Figure 3. RMSD of overlay of S protein RBD (pink = monkey and turquoise = human) complex with human (red) or monkey (blue) ACE2 after MD simulations showing different geometry of the two complexes.

provide evidence for or against infection of dogs[47]. Notably, our model predicted that companion animals would be permissive for SARS-CoV-2 infection, with this early prediction being supported by later published data on actual infection of companion animals[48,49].

Comparisons of our model predictions to other studies. Supplementary Fig. 5 summarizes the observed and predicted SARS-CoV-2 susceptibilities of species, inferred from analysis of phylogenetic clustering and sequence alignment of ACE2 of species known to be utilized by SARS-CoV-2 virus by Qiu et al.[50]. S protein-ACE2 interaction energies for various species were also predicted by Wu et al. using an automatic docking method, ICM-Pro[51]. However, these energies are inconsistent with the known susceptibilities of the relevant species, likely reflecting the fact that they were calculated based purely on docking calculations with no subsequent MD refinement. Rodrigues et al. published a study that used the HADDOCK docking method to estimate the relative strength of binding affinities of SARS-CoV-2 spike protein for ACE2 proteins of 30 species[52] with the docking including short restrained MD simulations. As noted by the authors, computational models have limitations, requiring validation against experimental data. For example, their model scored guinea pig and goldfish ACE2 among susceptible species, though it has been shown that guinea pig is non-susceptible[53], and indirect evidence suggests fish to be naturally resistant. This suggests that longer and potentially more accurate simulation protocols such as the one used in our work, are required to obtain sufficiently precise results to allow accurate comparison of differences in calculated binding energies. Other relevant modelling papers were also published while our paper was under review. Damas et al. published an analysis of ACE2 sequences from 410 vertebrate species, including 252 mammals, generating a probability that they could be used as a receptor by SARS-CoV-2[16]. They classed the species into five risk groups. Man, apes, monkeys, Chinese hamsters, whales and porpoises were included in the high-binding group. Golden hamsters, cattle and cats were members of the medium binding group while dogs, horses and bats were in the low binding group. Notably, pangolins, ferrets, mice, and minks were all assigned to the very low binding group in their analyses, susceptibility predictions that do not correlate well with either our data predictions or with actual experimental observations. Recently, another paper was published by Lam et al.[54] on the species specificity of S protein-ACE2 interaction. Like us, they used MODELLER to generate ACE2 structures for difference species and selected the best refined model using DOPE scores. They used free energy perturbation methods to calculate the binding energies of SARS-CoV-2 S protein to human and non-human ACE2 proteins although, unlike us, they did not include pangolin ACE2. Conspicuously, they generated energy differences similar to ours, this providing subsequent validation of our computational approach to calculating the relative binding affinities of non-human species. Two additional recent studies have relevance to our study. Spinello et al. published a detailed, microsecond MD simulation of the key molecular interactions driving the higher affinity of SARS-CoV-2 as compared to SARS to hACE2[33]. However, this study did not compare the affinities of S protein for non-human ACE2. Subsequently, Wang and co-workers published a MD study comparing the interactions of SARS-CoV-2 and SARS-CoV spike proteins with the human ACE2 protein with 200 ns simulations[34] but, again, no other species were considered.

Implications for the original SARS-CoV-2 animal source. Bats have been suggested as the original host species of SARS-CoV-2 infections in humans. Bat RaTG13 has the highest sequence similarity to SARS-CoV-2 with 96% whole-genome identity, but RaTG13 possesses neither the furin cleavage site nor the pangolin-CoV-like RBD seen in SARS-CoV-2 (Fig. 1, lower panel). Although bats carry many coronaviruses, no evidence of an immediate relative of SARS-CoV-2 in bat populations has been found so far. As highlighted by our data, the binding affinity of SARS-CoV-2 S protein for bat ACE2 is considerably lower than for human ACE2 and other species. Notably, RaTG13 S protein was shown not to bind to human ACE2[55]. Hence even if SARS-CoV-2 did originally arise from a bat precursor virus, which remains unproven, it must have spent considerable time in an intermediate animal host to allow it to adapt its S protein sufficiently to then be able to bind human ACE2. There are currently no explanations for how or where such a transition could have occurred to generate a SARS-CoV-2

spike protein optimised for human ACE2. Evidence of direct human infection by bat coronaviruses is rare, with transmission typically involving an intermediate host. For example, SARS CoV was found to be transmitted from bats to civet cats in which it first adapted before becoming able to infect humans. Hence SARS S protein had to acquire specific mutations to enable each species transition to occur, first to increase its affinity for civet ACE2 and then to increase its affinity for human ACE2. To date, a virus directly related to SARS-CoV-2 has not been identified in bats or any other non-human species, leaving its origins unclear. Wrobel et al. reported a structural biology study of bat RaTG13 and SARS-Cov-2 and calculated the binding affinities of human ACE2 for these viruses[55]. They concluded that, although the structures of SARS-CoV-2 and RaTG13 spike proteins are similar, SARS-CoV-2 spike has a more stable pre-cleavage form, with a k_D of 68 ± 9 nM for human ACE2 while RaTG13 bound human ACE2 almost 1000 times more weakly with a $k_D > 40$ μM. They observed that cleavage at the furin site decreased the overall stability of SARS-CoV-2 S protein and fostered the open conformation required for spike to bind to ACE2. They concluded that RaTG13 could not bind effectively to human ACE2 and hence would be unlikely to infect humans.

Early in the COVID-19 outbreak it was suggested that snakes may also be an intermediate vector for SARS-CoV-2. Turtle or snake ACE2 has very low homology to human ACE2 and our model predicted they would have minimal binding to S protein. Subsequently, SARS-CoV-2 was shown to not bind reptile ACE2, excluding snakes as an intermediate host for SARS-CoV-2[14]. As we discussed above, our model predicted pangolin ACE2 to have the closest binding affinity to hACE2, despite a low sequence similarity (86%). In contrast, monkey ACE2 with high similarity (97%) to hACE2 was predicted to have much lower binding affinity to S protein than human or pangolin ACE2. Binding affinities for all other modelled species were all substantially lower than for human ACE2 at the > 99.99% confidence level.

A coronavirus isolated from Malayan pangolins (pangolin-CoV) shared the same S protein RDB sequence as SARS-CoV-2, raising the suggestion that pangolins may have acted as an intermediate SARS-CoV-2 host between bats and humans. Although it has high spike RBD sequence similarity to SARS-CoV-2, pangolin-CoVs are not closely related to SARS-CoV-2, with ~ 90% sequence similarity across their whole genome[23]. It is noteworthy that the RBD common to both pangolin CoV and SARS-CoV-2 binds strongly to both pangolin and human ACE2, despite significant differences in these ACE2 molecules with only 63% of their binding site residues being common (Table 1), and the sequence similarity of ACE2 is only marginally higher between pangolins and human ACE2 ~ 85%, than between and bat and human ACE2 ~ 82%. Remarkably, Pangolin-CoV S protein has 100% amino acid identity with SARS-CoV-2 S protein but has much lower levels of identity of 98.6, 97.8 and 90.7% in the E, M and N proteins[24]. Pangolin CoVs isolated from Malayan pangolins from two different regions in China showed differences in the residues interacting with human ACE2[23]. One possibility might be that a pangolin was simultaneously co-infected with a bat ancestor to SARS-CoV-2 at the same time as being infected by a pangolin CoV. This could have allowed a recombination event to occur whereby the spike RBD of the pangolin CoV was inserted into the S protein of the bat CoV, thereby conferring the bat CoV with high binding for both pangolin and human ACE2. Such recombination events occur with other RNA viruses and explain creation of some pandemic influenza strains[56]. However, such events are rare as they require coinfection of the one host with two viruses at exactly the same time and the SARS-CoV-2 genome was reported to exhibit no evidence of recent recombination, arguing against this possibility[57]. Most importantly, if such a recombination event had occurred in pangolins it would be expected to have triggered an epidemic spread of the new highly permissive SARS-CoV-2-like virus among pangolin populations. Currently there is no evidence of a pangolin SARS-CoV-2 outbreak, making this scenario unlikely. Indeed, pangolins might be predicted to be protected from SARS-CoV-2 infection by the existence of cross-protective neutralising antibodies against pangolin coronaviruses given their close RBD similarity, making it even less likely that a SARS-CoV-2 was widely infecting pangolin populations and indeed no evidence of any such infection has been reported. Notably, all pangolin coronaviruses identified to date lack the furin-like cleavage site between S1/S2 in the SARS-CoV-2 S protein that facilitates its rapid spread through human populations. The fact pangolin CoV S protein does not have the furin cleavage site that is a prominent feature of the SARS-CoV-2 S protein[45], argues against pangolins being the intermediate vector for transmission of SARS-CoV-2 to humans. The major similarity of SARS-CoV-2 to pangolin-CoV lies in the S protein RBD residues that SARS-CoV-2 acquired by some unknown mechanism.

Does high human ACE2 binding affinity represent a recent gain-of-function mutation? Gain of function (GoF) mutations in viruses can lead to pandemics. For example, GoF mutations in influenza virus are associated with mammalian transmissibility, increased virulence for humans, and evasion of existing host immunity[56]. The progressive conditioning and adaptation of a new pandemic virus in humans is well recognized. Phylodynamic analyses of the COVID-19 genomes estimated the date for the most recent common ancestor as late 2019, consistent with the earliest reported date in 2019 for the initial cluster of pneumonia cases in Wuhan[58]. Based on available genome sequence data, this study concluded that the current pandemic has been driven entirely by human-to-human transmission since at least December. As the SARS-CoV-2 structure employed in our studies was obtained from viruses collected early in the outbreak, it is not clear how the very first SARS-CoV-2 strains acquired such a high affinity for human ACE2 without prior exposure. These data suggest that SARS-CoV-2 spike RBD evolved by selection on a human-like ACE2. While pangolin ACE2 has major differences in sequence and structure to hACE2, SARS-CoV-2 binding to pangolin ACE2 is second only to hACE2. Nevertheless, no SARS-CoV-2-like virus has been found in pangolins, suggesting they were not the original source. However, the fact pangolin CoVs can potentially use hACE2 for cell entry indicates that pangolin CoVs are a potential source of future human coronavirus pandemics, particularly if pangolin CoVs were to acquire the SARS-CoV-2 furin cleavage site.

Our study has several limitations, including the use homology models because ACE2 crystal structures for most species were not available, the use of the MMPBSA methods which tend to treat ligand entropy in an approximate manner, intrinsic limitations in the MD methods used (although we used an MD code that is widely recognized as robust), and the possibility that the furin cleavage site and different glycosylation patterns may influence the predicted binding affinities. We consider these issues are minimal as the ACE2 structures for most species are very similar and any deficiencies in the calculation of entropic contributions to the Gibbs free energy of binding and deficiencies in the MD force fields should largely cancel out. The propensity of the S protein RBD to adopt an open conformation is strongly influenced by the presence of the furin cleavage site and by glycosylation (e.g. Casalino et al.[59]). However, the S protein structure used for the binding energy comparisons is the same in all complexes with ACE2 proteins from different species, is derived from the experimental S protein-ACE2 complex structure, and thus is likely to result in cancellation of errors generated by these influences.

Given the seriousness of the ongoing SARS-CoV-2 pandemic, it is imperative that all efforts be made to identify the original source of the virus. It remains to be addressed whether SARS-CoV-2 is completely natural and was transmitted to humans by an intermediate animal vector or whether it might have arisen from a recombination event that occurred in a laboratory handling coronaviruses, inadvertently or intentionally, with the new virus being accidentally released into the local human population. Resolving these questions is of key importance so we can use such information to help prevent any similar outbreak in the future. In summary, our study suggests that from the beginning of this pandemic the SARS-CoV-2 S protein already had very high, optimal binding to hACE2. There is minimal early evidence of selection pressure to further optimise binding, in contrast to what has been seen with other zoonotic viruses at the time of their entry into the human population.

Materials and methods

Homology modelling of S protein and ACE2 from multiple species. As no three-dimensional structure of the SARS-CoV-2 S protein was available at the commencement of the project, we generated a homology structure the sequence retrieved from NCBI Genbank Database (accession number YP_009724390.1) in January 2020. A PSI-BLAST search against the PDB database for template selection was performed and the x-ray structure of SARS coronavirus S template (PDB ID 5XLR) was selected with 76.4% sequence similarity to SARS-CoV-2 S protein. The sequence alignment and sequences of related bat and pangolin coronaviruses are shown in Fig. 1.

The protein sequences of the ACE2 proteins for different species and full sequence alignment in are shown in Supplementary Fig. 2. The phylogenetic tree for ACE2 proteins from selected animal species is illustrated in Supplementary Fig. 1. Protein preparation and removal of non-essential and non-bridging water molecules for docking studies and analysis of docked proteins was performed using the UCSF Chimera package (https://www.cgl.ucsf.edu/chimera/).[60]

The 3D-structures of the RBD of SARS-Cov-2 S and non-human ACE2 proteins were built using Modeller 9.23 (https://salilab.org/modeller/),[61] using the SARS S protein template to generate the homology model of SARS-Cov-2 S. The ACE2 receptors of selected species were similarly homology modelled using the following template structures – 1R42 (human ACE2), 3CSI (human glutathione transferase) and 3D0G (ACE2 structure from spike protein receptor-binding domain from the 2002–2003 SARS coronavirus human strain complexed with human-civet chimeric receptor ACE2) (Supplementary Table 1). Template similarity is important for model building the model; the sequence of *Macaca fascicularis* (monkey, accession number A0A2K5X283) was 97% similar to that of human ACE2 while *Ophiophagus hannah* (king cobra) had a much lower similarity of 61% this template. The quality of the generated models was evaluated using the GA341 score[62] and DOPE ((Discrete Optimized Protein Energy) method scores[63], and the models assessed using SWISS-MODEL structure assessment server (https://swissmodel.expasy.org/assess).[64] Structures with the lowest DOPE score were refined by MD simulations (vide infra) and used for further analysis. We generated 10 homology models per protein that were refined and optimized in GROMACS.

The modelled structures were also assessed for quality control using Ramachandran Plot and molprobity scores in SWISSModel. The Ramachandran plot checks the stereochemical quality of a protein by analysing residue-by-residue geometry and overall structure geometry and visualizing energetically allowed regions for backbone dihedral angles ψ against φ of amino acid residues in protein structure. The Ramachandran score of SARS-CoV-2 spike protein was 90% in the binding region and the molprobity score was 3.17. The Ramachandran score of the percentage of amino acid residues in the various species ACE2 that fall into the energetically favoured region ranged from 96–99% (Supplementary Table 2). Notably, there were no Ramachandran outliers in the RBD. The predicted ACE2 structures and Ramachandran plots for each species are summarized in Supplementary Fig. 4. The molprobity provides evaluation of model quality at both the global and local level, this combines protein quality score that reflects the crystallographic resolution of a model[65]. It is a log-weighted combination of the number of serious atom clashes per 1000 atoms, percentage Ramachandran not favoured, and percentage bad side-chain rotamers. A good molprobity score is one that is equal to or lower than the crystallographic resolution. For reference, the MolProbity score for the x-ray structures of the templates (PDB IDs) were: 1R42 = 3.01; 3SCI = 3.14; 3D0G = 2.74 and 6M17 = 1.99. The Ramachandran and Molprobity scores show that all the built structures were of good quality and suitable for use in further studies.

The structure of the open form of the SARS-Cov-2 S protein was published subsequently (e.g., PDB ID 6VYB)[66]. Figure 4 shows the very high structural similarity of our homology modelled spike protein structure with the EM structures (PDB ID 6M0J (RBD) and 6VYB (open state)) with RMSD of 0.36 Å.

Docking of SARS-Cov-2 S protein with ACE2 proteins. These homology modelled ACE2 structures were docked against SARS-CoV-2 S protein structure using a state-of-the-art package HDOCK (http://hdock.

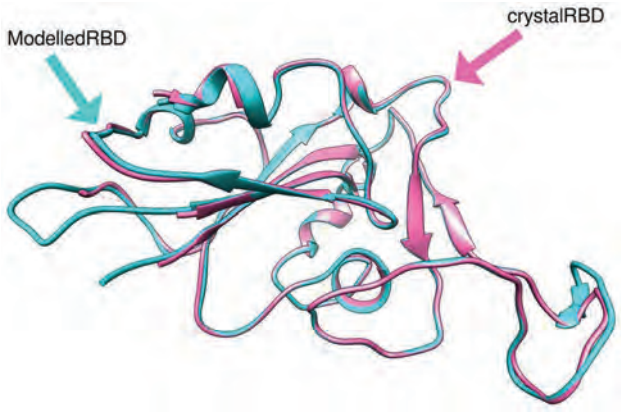


Figure 4. 3D structure of SARS-CoV-2 S protein (PDB ID 6M0J (open form) and the homology modelled structure from Modeller. This demonstrates the very high structural similarity of our homology modelled spike protein structure with the EM structures (PDB ID 6M0J (RBD)), with an RMSD of 0.36 Å.

phys.hust.edu.cn/).[67,68] This performs rigid-body docking by mapping the receptor and ligand molecules onto grids. It docks two molecules using an FFTW-based hierarchical approach. First, possible binding modes are globally sampled through an FFT-based global search strategy with an improved shape complimentary scoring method. Specifically, one molecule is fixed, and second molecule is rotated and translated in space. For each movement of the ligand, both the receptor and ligands molecules are mapped onto grids that extend past the proteins that account for long-range interactions of atoms. Molecular docking was performed on the homology modelled SARS-CoV-2 S protein and human and animal ACE2 proteins using the hybrid docking method because attempts to use the template-free docking method with the structures generated by Modeller were unsatisfactory. The hybrid method used the structural template for the complex (PDB ID 6M17) to generate the results reported here. All docking poses were ranked using an energy-based scoring function.

The hybrid docking procedure may potentially introduce bias into the structures of the S protein bound to ACE2 from non-human species because it uses a human ACE2 complex x-ray structure as a template. To check for possible bias, the ACE2 structures generated by HDOCK were compared with those generated independently by Modeller. The Cα backbones of the ACE2 structures aligned with RMSD values between 0.5 and 0.8 Å (see Supplementary Table 3), exhibiting very strong structural similarities. Additionally, complexes were subjected to molecular dynamics simulation to wash out any template-induced bias.

Molecular dynamics simulation of docked complexes. The final docked SARS-Cov-2 spike/ACE2 protein complexes were optimized using the AMBER99SB-ILDN force field in GROMACS2020 (http://www.gromacs.org/).[69] Simulations were carried out using the GPU accelerated version of the program and implementing periodic boundary conditions in ORACLE server. The final docked structures were selected by cluster analysis of the docked conformation and based on the RMSD analysis of docked conformation of our structures with 3D0G (SARS-RBD and ACE2). Docked complexes were immersed in a truncated octahedron box of TIP3P water molecules. The solvated box was further neutralized with Na + or Cl − counter ions using the tleap program. Particle Mesh Ewald (PME) was employed to calculate the long-range electrostatic interactions. The cut-off distance for the long-range van der Waals (VDW) energy term was 12.0 Å. The system was minimized without restraints. We applied 2500 cycles of steepest descent minimization followed by 5000 cycles of conjugate gradient minimization. After system optimization, the MD simulations was initiated by gradually heating each system in the NVT ensemble from 0 to 300 K for 50 ps using a Langevin thermostat with a coupling coefficient of 1.0/ps and with a force constant of 2.0 kcal/mol·Å2 on the complex. Finally, a production run of 100 ns of MD simulation was performed under a constant temperature of 300 K in the NPT ensemble with periodic boundary conditions for each system. During the MD procedure, the SHAKE algorithm was applied to all covalent bonds involving hydrogen atoms. The time step was 2 fs. The structural stability of the complex was monitored by the RMSD and RMSF values of the backbone atoms of the entire protein. Finally, the free energies of binding were calculated for all simulated docked structures.

Calculations were also performed for up to 500 ns to ensure that 100 ns is sufficiently long for convergence and that the docked conformation and protein–protein interaction was stable. We ran simulation of our docked spike RBD-human ACE2 for 500 ns and confirmed convergence by RMSD and RMSF. All complexes stabilized during simulations, with RMSD fluctuations converging to a range of 0.5 to 0.8 nm. We found that the complex had stabilised after 50 ns so considered that 100 ns was an adequate simulation time. The RMSD values for superimposition of the Cα backbones of each ACE2 structure before and after 100 ns simulation was 1.2 ± 0.1 Å, showing movement away from the initial HDOCK structures. We analysed the RMSF graph for the protein and did not observe much fluctuation in the amino acids. We used three production runs with different random starting seeds to estimate binding energies and binding energy uncertainties for each of the strongest binding ACE2 structures – human, bat, and pangolin. The binding energies in the Table 2 are based on a 10 ns analysis Sect. (1000 frames).

We also compared 100 ns simulated structures of the ACE2 proteins from all species against those generated by homology modelling (most x-ray structures were not available) and found the RMSD values for Cα alignments between 0.5 and 0.8 Å. This suggests that any memory of the human template has been removed or minimized. We also compared the structures generated independently by homology (Modeller) and HDOCK (Supplementary Table 3) and they agreed very well (RMSD < 1 Å).

Calculation of binding free energies of complexes. The binding free energies of the protein-protein complexes were evaluated in two ways. The traditional method is to calculate the energies of solvated SARS-CoV-2 S and ACE2 proteins and that of the bound complex proteins and derive the binding energy by subtraction.

$$\Delta G\left(\text{binding, aq}\right) = G\left(\text{complex, aq}\right) - G\left(\text{spike, aq}\right) - G\left(\text{ACE2, aq}\right) \quad (1)$$

We also calculated binding energies using the molecular mechanics Poisson Boltzmann surface area (MM-PBSA) tool in Gromacs that is derived from the nonbonded interaction energies of the complex[70,71]. The method is also widely used method for binding free energy calculations. The binding free energies of the protein complexes calculated using the most stable 10 ns section of the equilibrium phase from the output files of the 100 ns MD simulations. The `g_mmpbsa` tool in Gromacs was used after molecular dynamics simulations. It uses a number of MD snapshots to calculate free binding energy. We analysed the RMSD plot of all the complexes to understand the convergence of the MD trajectory. We identified the region where the complexes were most stable for 10 ns, then frames were extracted from this part of the trajectory to calculate the free binding energy using VdW, electrostatic, polar, and non-polar solvation energy. The output files obtained were used to post-process binding free energies by the single-trajectory MM-PBSA method. Specifically, for a non-covalent binding interaction in the aqueous phase the binding free energy,

$$\Delta G\left(\text{bind, aqu}\right) = \Delta G\left(\text{bind, vac}\right) + \Delta G\left(\text{bind, solv}\right) \quad (2)$$

where ΔG (bind,vac) is the binding free energy in vacuum, and ΔG(bind,solv) is the solvation free energy change upon binding:

$$\Delta G\left(\text{bind, solv}\right) = G\left(\text{R:L, solv}\right) - G\left(\text{R, solv}\right) - G\left(\text{L, solv}\right) \quad (3)$$

where G (R:L,solv), G (R,solv) and G (L,solv) are solvation free energies of complex, receptor and ligand, respectively[72].

Free energy decomposition analyses were also performed by MM-PBSA decomposition to get a detailed insight into the interactions between the ligand and each residue in the binding site. The binding interaction of each ligand–residue pair includes three terms: the van der Waals contribution, the electrostatic contribution, and the solvation contribution.

As the simulations are very lengthy, we only ran multiple simulations for the species proposed as intermediate hosts and potential sources of the original virus, human, pangolin, bat and snake to estimate the uncertainty in the binding energies. As the ACE2 proteins for all species were extremely similar in sequence, we expected that these simulation error estimates would be of the same order for all other species. The predicted binding energies for these other species were significantly lower than those of human ACE2. We also used a statistical test to calculate the probability of the pangolin and human ACE2 affinities being different.

Data availability

The coordinates of the S protein-ACE2 complexes will be deposited in. data repositories at La Trobe University and Flinders University.

Received: 14 August 2020; Accepted: 3 June 2021
Published online: 24 June 2021

References

1. Berkley, S. COVID-19 needs a big science approach. *Science* **367**, 1407. https://doi.org/10.1126/science.abb8654 (2020).
2. Sanders, J. M., Monogue, M. L., Jodlowski, T. Z. & Cutrell, J. B. Pharmacologic treatments for coronavirus disease 2019 (COVID-19): A review. *J. Am. Med. Assoc.* **323**, 1824–1836. https://doi.org/10.1001/jama.2020.6019 (2020).
3. Sohrabi, C. *et al.* World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *Int. J. Surg.* **76**, 71–76. https://doi.org/10.1016/j.ijsu.2020.02.034 (2020).
4. Le Thanh, T. *et al.* The COVID-19 vaccine development landscape. *Nat. Rev. Drug Discov.* **19**, 305–306. https://doi.org/10.1038/d41573-020-00073-5 (2020).
5. Almendros, A. & Gascoigne, E. Can companion animals become infected with COVID-19?. *Vet. Rec.* **186**, 419–420. https://doi.org/10.1136/vr.m1322 (2020).
6. Almendros, A. Can companion animals become infected with COVID-19?. *Vet. Rec.* **186**, 388–389. https://doi.org/10.1136/vr.m1194 (2020).
7. Ahmad, T. *et al.* COVID-19: Zoonotic aspects. *Travel Med. Infect. Dis.* **36**, 101607. https://doi.org/10.1016/j.tmaid.2020.101607 (2020).
8. Ashoor, D., Ben Khalaf, N., Marzouq, M., Jarjanazi, H. & Fathallah, M. D. SARS-CoV-2 RBD mutations, ACE2 genetic polymorphism, and stability of the virus-receptor complex: The COVID-19 host-pathogen nexus. *bioRxiv* https://doi.org/10.1101/2020.10.23.352344 (2020).
9. Kar, S. & Leszczynski, J. From animal to human: interspecies analysis provides a novel way of ascertaining and fighting COVID-19. *Innovation* **1**, 100021. https://doi.org/10.1016/j.xinn.2020.100021 (2020).
10. Lan, J. *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215–220. https://doi.org/10.1038/s41586-020-2180-5 (2020).

11. Abdel-Moneim, A. S. & Abdelwhab, E. M. Evidence for SARS-CoV-2 infection of animal hosts. *Pathogens* **9**, 529. https://doi.org/10.3390/pathogens9070529 (2020).
12. Katsnelson, A. How do viruses leap from animals to people and spark pandemics?. *Chem. Eng. News* **98**, 26 (2020).
13. Zhang, T., Wu, Q. & Zhang, Z. Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Curr. Biol.* **30**, 1346–1351. https://doi.org/10.1016/j.cub.2020.03.022 (2020).
14. Luan, J., Jin, X., Lu, Y. & Zhang, L. SARS-CoV-2 spike protein favors ACE2 from Bovidae and Cricetidae. *J. Med. Virol.* https://doi.org/10.1002/jmv.25817 (2020).
15. Taka, E. *et al.* Critical interactions between the SARS-CoV-2 spike glycoprotein and the human ACE2 receptor. *bioRxiv* https://doi.org/10.1101/2020.09.21.305490 (2020).
16. Damas, J. *et al.* Broad host range of SARS-CoV-2 predicted by comparative and structural analysis of ACE2 in vertebrates. *Proc. Natl. Acad. Sci. USA* **117**, 22311–22322. https://doi.org/10.1073/pnas.2010146117 (2020).
17. Luan, J., Lu, Y., Jin, X. & Zhang, L. Spike protein recognition of mammalian ACE2 predicts the host range and an optimized ACE2 for SARS-CoV-2 infection. *Biochem. Biophys. Res. Commun.* **526**, 165–169. https://doi.org/10.1016/j.bbrc.2020.03.047 (2020).
18. Piplani, S., Singh, P. K., Winkler, D. A. & Petrovsky, N. In silico comparison of spike protein-ACE2 binding affinities across species; significance for the possible origin of the SARS-CoV-2 virus. *arXiv* 2005.06199 (2020).
19. Rangel, H. R., Ortega, J. T., Maksoud, S., Pujol, F. H. & Serrano, M. L. SARS-CoV-2 host tropism: An in silico analysis of the main cellular factors. *Virus Res.* **289**, 198154. https://doi.org/10.1016/j.virusres.2020.198154 (2020).
20. Gimeno, A. *et al.* Prediction of novel inhibitors of the main protease (M-pro) of SARS-CoV-2 through consensus docking and drug reposition. *Int. J. Mol. Sci.* **21**, 3793 (2020).
21. Touyz, R. M., Li, H. & Delles, C. ACE2 the Janus-faced protein—From cardiovascular protection to severe acute respiratory syndrome-coronavirus and COVID-19. *Clin. Sci.* **134**, 747–750. https://doi.org/10.1042/CS20200363 (2020).
22. Shang, J. *et al.* Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**, 221–224. https://doi.org/10.1038/s41586-020-2179-y (2020).
23. Lam, T. T. *et al.* Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature* **583**, 282–285. https://doi.org/10.1038/s41586-020-2169-0 (2020).
24. Xiao, K. *et al.* Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature* **583**, 286–289. https://doi.org/10.1038/s41586-020-2313-x (2020).
25. Sit, T. H. C. *et al.* Infection of dogs with SARS-CoV-2. *Nature* **586**, 776–778. https://doi.org/10.1038/s41586-020-2334-5 (2020).
26. Opriessnig, T. & Huang, Y. W. Update on possible animal sources for COVID-19 in humans. *Xenotransplant.* **27**, e12621. https://doi.org/10.1111/xen.12621 (2020).
27. Chan, J. F. *et al.* Simulation of the clinical and pathological manifestations of Coronavirus Disease 2019 (COVID-19) in golden Syrian hamster model: Implications for disease pathogenesis and transmissibility. *Clin. Infect. Dis.* **71**, 2428–2446. https://doi.org/10.1093/cid/ciaa325 (2020).
28. Johansen, M. D. *et al.* Animal and translational models of SARS-CoV-2 infection and COVID-19. *Mucosal Immunol.* **13**, 877–891. https://doi.org/10.1038/s41385-020-00340-z (2020).
29. Shi, J. *et al.* Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS-coronavirus 2. *Science* **368**, 1016–1020. https://doi.org/10.1126/science.abb7015 (2020).
30. Kim, Y. I. *et al.* Infection and rapid transmission of SARS-CoV-2 in ferrets. *Cell Host Microbe* **27**, 704–709. https://doi.org/10.1016/j.chom.2020.03.023 (2020).
31. Segalés, J. *et al.* Detection of SARS-CoV-2 in a cat owned by a COVID-19–affected patient in Spain. *Proc. Natl. Acad. Sci. USA* **117**, 24790–24793. https://doi.org/10.1073/pnas.2010817117 (2020).
32. Yuan, M. *et al.* A highly conserved cryptic epitope in the receptor-binding domains of SARS-CoV-2 and SARS-CoV. *Science* **368**, 630–633. https://doi.org/10.1126/science.abb7269 (2020).
33. Spinello, A., Saltalamacchia, A. & Magistrato, A. Is the rigidity of SARS-CoV-2 spike receptor-binding motif the hallmark for its enhanced infectivity? Insights from all-atom simulations. *J. Phys. Chem. Lett.* **11**, 4785–4790. https://doi.org/10.1021/acs.jpclett.0c01148 (2020).
34. Wang, Y., Liu, M. & Gao, J. Enhanced receptor binding of SARS-CoV-2 through networks of hydrogen-bonding and hydrophobic interactions. *Proc. Natl. Acad. Sci. USA* **117**, 13967–13974. https://doi.org/10.1073/pnas.2008209117 (2020).
35. Ge, X. Y. *et al.* Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* **503**, 535–538. https://doi.org/10.1038/nature12711 (2013).
36. Li, W. *et al.* Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450–454. https://doi.org/10.1038/nature02145 (2003).
37. Wan, Y., Shang, J., Graham, R., Baric, R. S. & Li, F. Receptor recognition by the novel coronavirus from Wuhan: An analysis based on decade-long structural studies of SARS coronavirus. *J. Virol.* https://doi.org/10.1128/JVI.00127-20 (2020).
38. Korber, B. *et al.* Tracking changes in SARS-CoV-2 spike: Evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **182**, 812-827.e819. https://doi.org/10.1016/j.cell.2020.06.043 (2020).
39. Alexander, M. R. *et al.* Which animals are at risk? Predicting species susceptibility to COVID-19. *bioRxiv* https://doi.org/10.1101/2020.07.09.194563 (2020).
40. Mahdy, M. A. A., Younis, W. & Ewaida, Z. An overview of SARS-CoV-2 and animal infection. *Front. Vet. Sci.* **7**, 596391. https://doi.org/10.3389/fvets.2020.596391 (2020).
41. Temmam, S. *et al.* Absence of SARS-CoV-2 infection in cats and dogs in close contact with a cluster of COVID-19 patients in a veterinary campus. *One Health* **10**, 100164. https://doi.org/10.1016/j.onehlt.2020.100164 (2020).
42. Yu, P. *et al.* Age-related rhesus macaque models of COVID-19. *Animal Model Exp. Med.* **3**, 93–97. https://doi.org/10.1002/ame2.12108 (2020).
43. Rockx, B. *et al.* Comparative pathogenesis of COVID-19, MERS, and SARS in a nonhuman primate model. *Science* **368**, 1012–1015. https://doi.org/10.1126/science.abb7314 (2020).
44. Bao, L. *et al.* The pathogenicity of 2019 novel coronavirus in hACE2 transgenic mice. *Nature* **583**, 830–833 (2020).
45. Golden, J. W. *et al.* Human angiotensin-converting enzyme 2 transgenic mice infected with SARS-CoV-2 develop severe and fatal respiratory disease. *JCI Insight* **5**, e142032. https://doi.org/10.1172/jci.insight.142032 (2020).
46. Huang, J., Song, W., Huang, H. & Sun, Q. Pharmacological therapeutics targeting RNA-dependent RNA polymerase, proteinase and spike protein: from mechanistic studies to clinical trials for COVID-19. *J. Clin. Med.* **9**, 1131 (2020).
47. Goumenou, M., Spandidos, D. A. & Tsatsakis, A. [Editorial] Possibility of transmission through dogs being a contributing factor to the extreme Covid19 outbreak in North Italy. *Mol. Med. Rep.* **21**, 2293–2295. https://doi.org/10.3892/mmr.2020.11037 (2020).
48. Ferasin, L. *et al.* Myocarditis in naturally infected pets with the British variant of COVID-19. *bioRxiv* https://doi.org/10.1101/2021.03.18.435945 (2021).
49. Grimm, D. Major coronavirus variant found in pets for first time. *Science* https://doi.org/10.1126/science.abi6152 (2021).
50. Qiu, Y. *et al.* Predicting the angiotensin converting enzyme 2 (ACE2) utilizing capability as the receptor of SARS-CoV-2. *Microbes Infect.* **22**, 221–225. https://doi.org/10.1016/j.micinf.2020.03.003 (2020).
51. Wu, C. *et al.* In silico analysis of intermediate hosts and susceptible animals of SARS-CoV-2. *ChemRxiv* https://doi.org/10.26434/chemrxiv.12057996.v1 (2020).

52. Rodrigues, J. *et al.* Insights on cross-species transmission of SARS-CoV-2 from structural modeling. *PLoS Comput. Biol.* **16**, e1008449. https://doi.org/10.1371/journal.pcbi.1008449 (2020).
53. Li, Y. *et al.* SARS-CoV-2 and three related coronaviruses utilize multiple ACE2 orthologs and are potently blocked by an improved ACE2-Ig. *J. Virol.* **94**, e01283-e11220. https://doi.org/10.1128/JVI.01283-20 (2020).
54. Lam, S. D. *et al.* SARS-CoV-2 spike protein predicted to form complexes with host receptor protein orthologues from a broad range of mammals. *Sci. Rep.* **10**, 16471. https://doi.org/10.1038/s41598-020-71936-5 (2020).
55. Wrobel, A. G. *et al.* SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. *Nat. Struct. Mol. Biol.* **27**, 763–767. https://doi.org/10.1038/s41594-020-0468-7 (2020).
56. Casadevall, A. & Imperiale, M. J. Risks and benefits of gain-of-function experiments with pathogens of pandemic potential, such as influenza virus: a call for a science-based discussion. *MBio* **5**, e01730-e1714. https://doi.org/10.1128/mBio.01730-14 (2014).
57. Day, T., Gandon, S., Lion, S. & Otto, S. P. On the evolutionary epidemiology of SARS-CoV-2. *Curr. Biol.* **30**, R849–R857. https://doi.org/10.1016/j.cub.2020.06.031 (2020).
58. Rambaut, A. in *Virological.org* (2020).
59. Casalino, L. *et al.* Beyond shielding: The roles of glycans in the SARS-CoV-2 spike protein. *ACS Cent. Sci.* **6**, 1722–1734. https://doi.org/10.1021/acscentsci.0c01056 (2020).
60. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612. https://doi.org/10.1002/jcc.20084 (2004).
61. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815. https://doi.org/10.1006/jmbi.1993.1626 (1993).
62. John, B. & Sali, A. Comparative protein structure modeling by iterative alignment, model building and model assessment. *Nucleic Acids Res.* **31**, 3982–3992. https://doi.org/10.1093/nar/gkg460 (2003).
63. Shen, M. Y. & Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **15**, 2507–2524. https://doi.org/10.1110/ps.062416606 (2006).
64. Benkert, P., Biasini, M. & Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinform.* **27**, 343–350. https://doi.org/10.1093/bioinformatics/btq662 (2011).
65. Chen, V. B. *et al.* MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21. https://doi.org/10.1107/S0907444909042073 (2010).
66. Walls, A. C. *et al.* Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **181**, 281–292. https://doi.org/10.1016/j.cell.2020.02.058 (2020).
67. Yan, Y., Zhang, D., Zhou, P., Li, B. & Huang, S. Y. HDOCK: A web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy. *Nucl. Acids Res* **45**, W365–W373. https://doi.org/10.1093/nar/gkx407 (2017).
68. Yan, Y., Tao, H., He, J. & Huang, S. Y. The HDOCK server for integrated protein-protein docking. *Nat. Protoc.* **15**, 1829–1852. https://doi.org/10.1038/s41596-020-0312-x (2020).
69. Abraham, M. J. *et al.* GROMACS; High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
70. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA* **98**, 10037–10041. https://doi.org/10.1073/pnas.181342398 (2001).
71. Kumari, R., Kumar, R., Open Source Drug Discovery, C. & Lynn, A. g_mmpbsa—A GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* **54**, 1951–1962, https://doi.org/10.1021/ci500020m (2014).
72. Wang, E. *et al.* End-point binding free energy calculation with MM/PBSA and MM/GBSA: Strategies and applications in drug design. *Chem. Rev.* **119**, 9478–9508. https://doi.org/10.1021/acs.chemrev.9b00055 (2019).

Acknowledgements

We would like to thank Harinda Rajapaksha for assistance to optimise GROMACS for this project. We would also like to thank Oracle for providing their Cloud computing resources for the modelling studies described herein. In particular, we wish to thank Peter Winn, Dennis Ward, and Alison Derbenwick Miller from Oracle in facilitating these studies. The opinions expressed herein are solely those of the individual authors and should not be inferred to reflect the views of their affiliated institutions, funding bodies or Oracle corporation.

Author contributions

N.P.—conceived project, analysed data, contributed to manuscript; S.P. and P.K.S.—performed the computations, analysed data, contributed to the manuscript; D.A.W.—analysed data and contributed to manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-92388-5.

Correspondence and requests for materials should be addressed to D.A.W. or N.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Exhibit

34

NEWS IN FOCUS

WUHAN VIROLOGY INSTITUTE

Hazard suits hang at the National Bio-safety Laboratory, Wuhan, the first lab on the Chinese mainland equipped for the highest level of biocontainment.

INFECTIOUS DISEASE

Inside China's pathogen lab

Maximum-security biosafety facility nears approval, sparking excitement and concern.

BY DAVID CYRANOSKI, WUHAN, CHINA

A laboratory in Wuhan is on the cusp of being cleared to work with the world's most dangerous pathogens. The move is part of a plan to build between five and seven biosafety level-4 (BSL-4) labs across the Chinese mainland by 2025, and has generated much excitement, as well as some concerns.

Some scientists outside China worry about pathogens escaping, and the addition of a biological dimension to geopolitical tensions between China and other nations. But Chinese microbiologists are celebrating their entrance to the elite cadre empowered to wrestle with the world's greatest biological threats.

"It will offer more opportunities for Chinese researchers, and our contribution on the BSL-4-level pathogens will benefit the world," says George Gao, director of the Chinese Academy of Sciences Key Laboratory of Pathogenic Microbiology and Immunology in Beijing. There are already two BSL-4 labs in Taiwan, but the National Bio-safety Laboratory, Wuhan, would be the first on the Chinese mainland.

The lab was certified as meeting the standards and criteria of BSL-4 by the China National Accreditation Service for Conformity Assessment (CNAS) in January. The CNAS examined the lab's infrastructure, equipment and management, says a CNAS representative, paving the way for the Ministry of Health to give its approval. A representative from the ministry says it will move slowly and cautiously; if the assessment goes smoothly, it could approve the laboratory by the end of June.

BSL-4 is the highest level of biocontainment: its criteria include filtering air and treating ▶

MUYI XIAO FOR *NATURE*

The central monitor room at China's National Bio-safety Laboratory.

water and waste before they leave the laboratory, and stipulating that researchers change clothes and shower before and after using lab facilities. Such labs are often controversial. The first BSL-4 lab in Japan was built in 1981, but operated with lower-risk pathogens until 2015, when safety concerns were finally overcome.

The expansion of BSL-4-lab networks in the United States and Europe over the past 15 years — with more than a dozen now in operation or under construction in each region — also met with resistance, including questions about the need for so many facilities.

The Wuhan lab cost 300 million yuan (US$44 million), and to allay safety concerns it was built far above the flood plain and with the capacity to withstand a magnitude-7 earthquake, although the area has no history of strong earthquakes. It will focus on the control of emerging diseases, store purified viruses and act as a World Health Organization 'reference laboratory' linked to similar labs around the world. "It will be a key node in the global biosafety-lab network," says lab director Yuan Zhiming.

The Chinese Academy of Sciences approved the construction of a BSL-4 laboratory in 2003, and the epidemic of SARS (severe acute respiratory syndrome) around the same time lent the project momentum. The lab was designed and constructed with French assistance as part of a 2004 cooperative agreement on the prevention and control of emerging infectious diseases. But the complexity of the project, China's lack of experience, difficulty in maintaining funding and long government approval procedures meant that construction wasn't finished until the end of 2014.

The lab's first project will be to study the BSL-3 pathogen that causes Crimean–Congo haemorrhagic fever: a deadly tick-borne virus that affects livestock across the world, including in northwest China, and that can jump to people.

Future plans include studying the pathogen that causes SARS, which also doesn't require a BSL-4 lab, before moving on to Ebola and the West African Lassa virus, which do. Some one million Chinese people work in Africa; the country needs to be ready for any eventuality, says Yuan. "Viruses don't know borders."

Gao travelled to Sierra Leone during the recent Ebola outbreak, allowing his team to report the speed with which the virus mutated into new strains (Y.-G. Tong *et al. Nature* **524,** 93–96; 2015). The Wuhan lab will give his group a chance to study how such viruses cause disease, and to develop treatments based on antibodies and small molecules, he says.

"Viruses don't know borders."

The opportunities for international collaboration, meanwhile, will aid the genetic analysis and epidemiology of emergent diseases. "The world is facing more new emerging viruses, and we need more contribution from China," says Gao. In particular, the emergence of zoonotic viruses — those that jump to humans from animals, such as SARS or Ebola — is a concern, says Bruno Lina, director of the VirPath virology lab in Lyon, France.

Many staff from the Wuhan lab have been training at a BSL-4 lab in Lyon, which some scientists find reassuring. And the facility has already carried out a test-run using a low-risk virus.

But worries surround the Chinese lab, too. The SARS virus has escaped from high-level containment facilities in Beijing multiple times, notes Richard Ebright, a molecular biologist at Rutgers University in Piscataway, New Jersey. Tim Trevan, founder of CHROME Biosafety and Biosecurity Consulting in Damascus, Maryland, says that an open culture is important to keeping BSL-4 labs safe, and he questions how easy this will be in China, where society emphasizes hierarchy. "Diversity of viewpoint, flat structures where everyone feels free to speak up and openness of information are important," he says.

Yuan says that he has worked to address this issue with staff. "We tell them the most important thing is that they report what they have or haven't done," he says. And the lab's international collaborations will increase openness. "Transparency is the basis of the lab," he adds.

The plan to expand into a network heightens such concerns. One BSL-4 lab in Harbin is already awaiting accreditation; the next two are expected to be in Beijing and Kunming, the latter focused on using monkey models to study disease.

Lina says that China's size justifies this scale, and that the opportunity to combine BSL-4 research with an abundance of research monkeys — Chinese researchers face less red tape than those in the West when it comes to research on primates — could be powerful. "If you want to test vaccines or antivirals, you need a non-human primate model," says Lina.

But Ebright is not convinced of the need for more than one BSL-4 lab in mainland China. He suspects that the expansion there is a reaction to the networks in the United States and Europe, which he says are also unwarranted. He adds that governments will assume that such excess capacity is for the potential development of bioweapons.

"These facilities are inherently dual use," he says. The prospect of ramping up opportunities to inject monkeys with pathogens also worries, rather than excites, him: "They can run, they can scratch, they can bite."

Trevan says China's investment in a BSL-4 lab may, above all, be a way to prove to the world that the nation is competitive. "It is a big status symbol in biology," he says, "whether it's a need or not." ■

UPDATE

The name of the CNAS representative has been removed from this article on request.

Exhibit

35

Is 'Patient Su' Covid's Patient Zero? IAN BIRRELL, who's led the way in exposing Beijing's lies, reveals how a woman aged 61 was diagnosed with virus THREE weeks before China admits that it even existed

- **Mistake by Chinese official may have disclosed details of early Covid case**
- **Error shows the woman, 61, lived about a mile from a coronavirus research lab**
- **She was also close to a stop for the high-speed rail line that spread the virus**

By IAN BIRRELL FOR THE MAIL ON SUNDAY
PUBLISHED: 17:06 EDT, 29 May 2021 | **UPDATED:** 03:20 EDT, 30 May 2021

A mistake by a leading Chinese official may have disclosed the name, address and details about one of the first people suspected of being infected with Covid-19 in Wuhan, three weeks before **Beijing** authorities claim they detected the initial case.

The astonishing error, revealed in a screen-grab sent to a Chinese medical journal, shows that the 61-year-old woman, known as 'Patient Su', lived about a mile from one of the city's main **coronavirus** research labs.

She was also close to a stop for the high-speed rail line that is believed to have played a key role in spreading the virus around the city of 11 million people.

Tom Tugendhat MP, chairman of the Common's foreign affairs committee, said: 'The time has come for **China** to open up all its files so the world can find the truth about the origins of this pandemic.

'We cannot protect against future risks if there is not recognition that we all need to share knowledge and learn from any mistakes.'

The astonishing error, revealed in a screen-grab sent to a Chinese medical journal, shows that the 61-year-old woman, known as 'Patient Su', lived about a mile from one of the city's main coronavirus research labs (pictured)

This latest development emerged as the result of an interview given to a Chinese medical journal by the scientist tasked with compiling the country's official data on cases.

Professor Yu Chuanhua, professor of biostatistics at Wuhan University, told Health Times that he had 47,000 cases on his national database of confirmed and suspected cases by late February 2020.

These included one suspected fatality of a patient who fell ill in late September 2019.

'There is data on a patient who became ill on September 29,' he said. 'The data shows the patient has not undergone nucleic testing and the clinical diagnosis is a suspected case. The patient has died. The data has not been confirmed.'

The academic then detailed two more suspected cases reported to Wuhan doctors on November 14 and 21, along with several others before December 8 – the date that China gave to the World Health Organisation for the 'earliest onset case'.

The Health Times article included a screenshot of the two November cases on the professor's database. Although personal details were blurred out, some were visible, including the hospital name and home district.

They show Patient Su was treated at Rongjun Hospital in Wuhan and, given the building and street numbers, almost certainly lived in the Kaile Guiyan community on Zhuodaoquan Street, about 600 metres from the medical centre.

Both the hospital and the residence are in the Hongshan district near where much of the bat-related coronavirus research was taking place in several laboratories.

This latest development emerged as the result of an interview given to a Chinese medical journal by the scientist tasked with compiling the country's official data on cases. Pictured, the Wuhan Institute of Virology

These include a laboratory run by China's Centre for Disease Control with the second-highest global levels of biosecurity little more than one mile away, while downtown sites run by Wuhan Institute of Virology are less than three miles away.

David Asher, former lead investigator for the US State Department, told The Mail on Sunday in March that three researchers at the institute had become ill with a mysterious respiratory condition in November 2019, with the wife of one scientist dying.

TikTok star's stepdad tries to kill her in attempted murder suicide

8.8k viewing now

Heartbroken elderly woman, 96, evicted from home in Dem state

2.8k viewing now

Iconic national park abruptly cancels all summer vacations

4.1k viewing now

Last week, the Wall Street Journal reported these researchers ended up in hospital although China has furiously denied the claims. On Thursday, President Joe Biden ordered US intelligence services to carry out a fresh probe into the pandemic's origins.

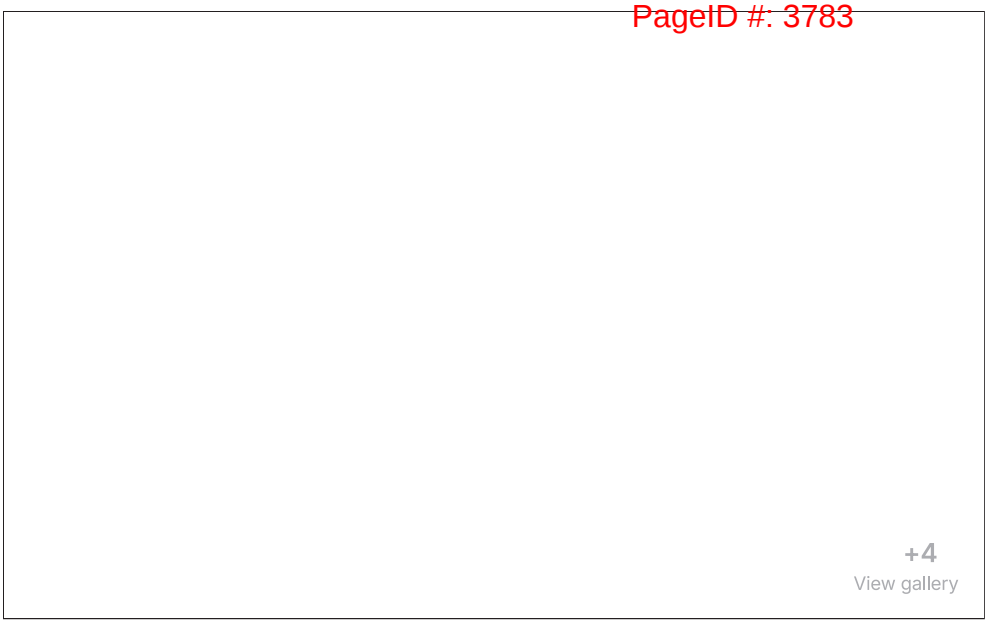
One Washington source told The Mail on Sunday that US intelligence on the Wuhan researchers was collected in late 2019 in data-scraping from routine surveillance. It is thought to include tapped phone conversations, texts and emails.

He said it was not discovered until efforts were intensified last year to investigate the pandemic's origins and any possible links with Wuhan laboratories – and that it is backed by testimony from a source with access to one of the units.

It is also understood that Washington has been surprised by the lack of input from British intelligence into China's cover-up, given their liaison over other global concerns. 'The UK seemed surprisingly reticent,' he said.

The area where Patient Su lived and was treated is more than 13 miles from the Huanan market originally blamed by Beijing as the source of Covid-19 and which was rapidly cleaned up after Taiwan notified the WHO about the emerging crisis. The second patient in November was listed as a 62-year-old man called Wang, who was treated at Hanyang hospital.

Professor's Yu's interview with Health Times took place on the day China's health authorities issued a silencing gag on the novel coronavirus as President Xi Jinping tried to regain control of the situation.



The area where Patient Su lived and was treated is more than 13 miles from the Huanan market originally blamed by Beijing as the source of Covid-19 and which was rapidly cleaned up after Taiwan notified the WHO about the emerging crisis. Pictured, the Wuhan Institute of Virology

Yu rang the journalist within two days to retract this information, claiming the dates had been entered incorrectly and all the other suspected cases before December 8 needed verification.

The details were discovered by Gilles Demaneuf, a member of the 'Drastic' group of online digital activists who have uncovered many of the facts seen as contradicting the official Chinese narrative that Covid-19 was a disease that crossed over naturally from animals.

'We were able to pinpoint the exact name, age and address of a very early suspected case nearly one month before the official first case,' said Demaneuf, a French data scientist who works for a New Zealand bank. 'That address is right next to the subway line No 2 and also not far from a People's Liberation Army hospital that treated some of the other earliest cases.'

This rail system carries a million people a day and connects the wet market, Wuhan Institute of Virology and an international airport.

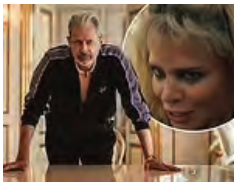
Demaneuf argues these new findings show many more clues might be accessible if there are continued and determined efforts to evaluate the lab leak theory – rather than 'wishful acceptance at face value of statements from China'. The WHO was widely condemned for its 'whitewash' investigation earlier this year when it allowed Beijing to vet its team of experts.

It adopted China's narrative that a lab leak was 'extremely unlikely' while pushing discredited theories that Covid might have been imported on frozen food. The organisation's joint report even backtracked on an influential Lancet study by Chinese scientists which examined the first 41 patients admitted to hospital and gave the date of the first case as December 1, 2019.

This man, whose wife and son also fell ill as the first confirmed family cluster of Covid, was excluded because his respiratory condition allegedly responded to antibiotics.

Meanwhile, a well-sourced report in the South China Morning Post in March claimed there were nine cases by the end of November, involving four men and five women

J.D. Vance shuts down 'haters' who BOOED him during firefighters speech by revealing his initial thoughts on Donald Trump


Netflix fans 'completely obsessed' with 'glorious' new show and demand a second season after binge-watching - branding creator 'a genius'


American woman, 18, is arrested in Greece for 'raping 20-year-old drunk man in an abandoned house'


Bone-chilling home surveillance footage of prowler in elderly woman's house that would terrify anyone


Bristol Palin reveals 'gut-wrenching' six-word text teen son sent her about moving to live with dad and stepmom 3,000 miles away


MSNBC anchor goes haywire on Trump advisor before ending segment with furious threat


California parents terrified after group of 20 migrants try to board elementary school bus


How this tragic card about becoming a mother sealed a young girl's death warrant after years of shocking abuse


aged between 39 and 79, with the first diagnosis on November 17.

Southampton University modelling experts suggested China could have cut global cases by 95 per cent if it had taken action to contain the disease three weeks earlier – instead of covering up the outbreak and pressing ahead with New Year festivities that involved millions of people moving around the country.

President Biden will release origins of COVID intelligence report

Watch the full video

0:00 / 1:04

Beijing China Covid-19

Share or comment on this article: Is 'Patient Su' Covid's Patient Zero? asks IAN BIRRELL

278 shares

Exhibit

36

palgrave communications
HUMANITIES | SOCIAL SCIENCES | BUSINESS

ARTICLE

Received 4 Dec 2015 | Accepted 6 Jun 2016 | Published 12 Jul 2016

DOI: 10.1057/palcomms.2016.45

OPEN

Scientific advice in China: the changing role of the Chinese Academy of Sciences

Xiaoxuan Li[1], Kejia Yang[1] and Xiaoxi Xiao[1]

ABSTRACT This article explores the changing role of the Chinese Academy of Sciences (CAS) in the national scientific advisory system in China from an historical perspective, which has been divided into three periods. Period I (from 1949 to 1976) relates to the era during which, since its establishment, CAS assumed the role of the top-level scientific advisor to the Chinese government, and its dual role of "scientific advisory body" and "science and technology administrative organ" were especially evident during the formulation of "The Twelve-Year National Long-term Outline for Science and Technology Development (1956–1967)". During period II (from 1977 to 2011) the function of providing scientific advice to the government was restored after the "Cultural Revolution", and CAS entered a new era of unprecedented prosperity and development. Finally, period III (from 2012 to present) includes Chinese President Xi Jinping taking office, and the central government encouraging the development of science and technology think tanks, using third-party evaluation as an important measure and approach to promote the reform of the government management system. In summary, the new government attaches great importance to CAS, urging it, as a third-party organization, to fully play its role as a scientific advisor and further develop its capability and approaches to consulting, and improve its organizational structure. Given its prominent position in the Chinese scientific advisory system, the changing role of CAS also reflects the characteristics and changes in China's scientific advice throughout different periods. This article is published as part of a collection on scientific advice to governments.

[1] Institute of Policy and Management, Chinese Academy of Sciences, Beijing, China

Introduction

Scientific advice has been a significant factor to be considered by national governments of various countries when making policies (Jasanoff, 2009; Arimoto and Sato, 2012; Australian and Government Department of Industry, 2012; Doubleday and Wilsdon, 2013). However, given the differences in political institutions and cultures across countries, the arrangements and methods for scientific advice are various, which have been widely discussed among scholars (Glynn *et al.*, 2002; Wilsdon *et al.*, 2014; OECD, 2015). National academies, as important resources for providing scientific advice to governments, have played a significant role in the national system, such as in the United States, the United Kingdom and other countries (Blair, 2006; Collins, 2011).

The Chinese Academy of Sciences (CAS), founded in 1949, is a primary scientific advisory institution in China (Cao, 2004). It comprises three major parts, which are a merit-based learned society, a comprehensive research and development network and a system of higher education.[1] The merit-based part of Chinese Academy of Sciences is represented by the Academic Divisions of the Chinese Academy of Sciences (CASAD), with more than 700 members and 70 plus foreign members. The comprehensive research and development part of CAS consists of over 100 research institutions, with research areas covering almost the whole spectrum of the natural sciences. It includes 60,000 plus researchers and its funding for scientific research reached 45 billion RMB in 2014. Furthermore, there are more than 50,000 students studying in CAS. As a top research institution and national academy, CAS has held an important position in the Chinese advisory system since its establishment, and plays a significant role in providing scientific advice to government especially on those complicated and comprehensive issues involving science and technology.

Recently, the Chinese government has attached growing importance to the construction of think tanks, and particularly emphasizes the role of science and technology think tanks in government's advisory system. In 2013, Chinese President Xi Jinping urged CAS not only to play a leading role in scientific research but also build a high-caliber national science and technology think tank during his visit to CAS. Although there have been many reports about CAS's academic research and development strategies, demonstrating its role as a research institution, only a few introductions have been made to show its role as a scientific advisory body. Cao and Suttmeier (1999) introduce the policy roles of CAS members, and especially focus on the role of academicians, in providing recommendations for central government before 2000. However, this article explores the changing role of CAS, including both its academicians and research institutions, in Chinese scientific advisory system from an historical perspective, which can be divided into three periods. It is worth pointing out that the role of CAS in different periods reflects the characteristics and changes of the modes and approaches of Chinese scientific advice as for CAS's prominent position in the system.

The highest-level scientific advisory body to government since its establishment

Period I (1949–1976). After the People's Republic of China was founded, the Chinese government gathered almost all of the outstanding domestic scientists to establish CAS. In the early time of its establishment, apart from playing its role as a center of national scientific research, CAS performed as the administrative organ of national science and technology[2](Qian, 2010). As one of the government departments affiliated to the Government Administration Council of the Central People's Government (latter was named the State Council), CAS could make its own policies and decisions and also could coordinate all of the science and technology resources across the country to formulate and deliver policies.

As China's largest research body and also the important national research institution, CAS plays a significant role in providing scientific advice to support national science and technology development, which could be highlighted in the formulation of "The Twelve-Year National Long-term Outline for Science and Technology Development (1956–1967) (i.e., The Twelve-Year Plan (1956–1967))" with the dual role of "scientific advisory body" and "science and technology administrative organ".

At the early stage of 1950s, China entered the first stage of the planned economic system. To meet the needs of economic growth, it is necessary for China to have a long-term plan for the development of science and technology. In 1953, the CAS delegation of 26 scientists visited the Soviet Union to learn its experience on the development of science and technology. During this visit, it is one of the important concern of Chinese scientists that how to formulate the development plan of science and technology. It is after this visit that the Chinese scientists begin to consider the formation of academic divisions in CAS.

In 1955, the CASAD was founded as the leading body of the academia and the most authoritative academic senate in China. The formulation of the 15 years long-term development plan of science and technology was advocated at the inaugural meeting of the foundation of CASAD in June. The aim of this plan was to fully play the role of science and technology in the development of society and economy, and to meet the need of the national long-term development plan. In January 1956, En-lai Chou advocated the State Planning Committee and other related government departments working out the "The Twelve-Year National Long-term Outline for Science and Technology Development (1956–1967)" during the report on the question of intellectuals. Then the State Planning Committee of Science with 35 members was founded, led by Chen Yi, to be in charge of[3] the working out of "The Twelve-Year Plan" (Wu, 1994).

"The Twelve-Year Plan" is the first long-term plan to encourage the development of science and technology in New China. This plan set out the general objectives of the national science and technology development, putting forward 57 programs and 600 plus research subjects, involving important fields and crucial issues of scientific and technological development, such as semiconductor technology, computer technology, automation technology, radio technology, nuclear technology and jet technology(Li, 1991; Hu, 2006; Yang and Zhang, 2007). Thanks to this plan, a large number of research institutions were established, and the number of research staff also had a substantial increase, which laid the foundation for the establishment of New China's science and technology system promoting the emergence and development of a series of new industries and sectors of industries in agriculture and national defense. This plan was formulated through the support of entire nation, and the CASAD played a leading role to coordinate more than 600 scientists and engineers under the leadership of the State Planning Committee of Science(Li, 1999).

There are two major stages to work out this plan. *Stage 1*: Drafts of departmental plan were submitted by CAS, universities, industries and national defense departments respectively at the end of February 1956. *Stage 2*: Including the academicians in the Division of Mathematics & Physics & Chemistry, the Division of Life Sciences & Earth Sciences and the Division of Technological Sciences in CAS, the CASAD organized more than 600 scientists and engineers, working in a general group and several special groups, to collect, integrate and review the separated draft plans

into a synthesis report. In August 1956, The Twelve-Year Plan was completed through the underpinning efforts of CAS.

As the administrative function of CAS had seriously affected its construction as national scientific research center, CAS had no longer undertaken the function of national science and technology administration since 1950s, which was transferred to the State Science and Technology Commission[4] founded in 1958, and CAS became one of the highest academic institutions under the leadership of the State Council. At the same time, the advisory function of CAS was also partly replaced by the State Science and Technology Commission, such as the role of working out the long-term plan for the development of national science and technology.

Apart from organizing and coordinating the scientists and engineers across the country to work out The Twelve-Year Plan, CAS also provided some other important scientific advices as the top scientific advisory institution until the occurrence of Cultural Revolution in 1966(Wang, 2015). In the 10-year-long turbulence, the scientific advisory work of CAS was seriously affected[5] and the General Assembly of the CAS Members was suspended for 20 years after the Third Assembly held in 1960.

During this period (1949–1976), science was easily affected by the politics especially in the context of the planned economic system. Scientific advices were often used as explanations and descriptions of the political slogans rather than the objective and sound evidences to support decisions. As the top institution, CAS played a significant role as the highest scientific advisory body to support the strategy development of the national science and technology especially at the birth of New China, where a thousand things wait to be done. However, due to the incomplete scientific advisory system, CAS was easily affected by politics, especially during the Cultural Revolution, the survival of scientists were seen as problems, let alone to guarantee the freedom of providing scientific opinions (Cao and Suttmeier, 1999).

The restoration and prosperity of CAS's function as a scientific advisor

Period II (1977–2011). After Cultural Revolution, the government increasingly emphasizes the importance of science and technology in serving economic growth and social progress, advocating the strategic concept of "science and technology are the primary productive force", "Economic development must rely on science and technology, and the latter should be oriented toward the former", "Science and Education for a Prosperous China"(Li *et al.*, 2015). Meanwhile, the Chinese government vigorously promoted the scientific and democratic construction of the government's decision making system (Wang, 2008; Hou, 2009), thus giving an extended period of scientific advice prosperity in China(Cheng, 2005). During this period, scientific advice not only played an important role in policy-making for development of science and technology, but also actively exerted influences on policy-making on other science and technology issues related to social and economic growth(Cao, 2004; Zheng *et al.*, 2013).

In January 1979, after the national science conference was held, the activities of CASAD was restored, and 2 years later the Fourth General Assembly of CAS Members was convened,[6] now the Assembly convenes regularly every 2 years(Wei, 2006). During this booming period, both Academic Divisions and research institutions of CAS leveraged its advantages and strengthened the cooperative communication with other departments, so as to give a full play to its role as a scientific advisor to government's decision making through a variety of channels. As the highest national advisory body on science and technology and a traditional merit-based academic society, the CASAD provides advice mainly through academicians. CAS's research institutions have large research resources to support scientific advisory work, by taking advantage of different strengths of research institutes, such as organizing interdisciplinary specialists and scholars across institutes.

Providing scientific advices to government mainly through CAS academicians. In 1984, the State Council, China's cabinet, officially stated that the CASAD is the highest national advisory body on science and technology. Since then, the CASAD had gradually attached more attention on providing scientific advice through CAS academicians rather than playing the role of an academic senate as in its initial stage (Wang, 2015). To improve the advisory role of CASAD, the government allocated funding to support academicians to conduct advisory activities, and more internal and external meetings were held to discuss research priorities and policy issues. For example, the General Assembly of the Members of CAS and CAE (Chinese Academy of Engineering) is held biennially, and top leaders of the central government attend the meetings, discussing the important subjects of scientific research in the next few years and how the development of science could fully serve national needs.

The consulting activities of CAS academicians mainly focus on the important fields of national long-term interests, for instance, the long-term science development plans and large-scale programs, and crucial and complicated issues of economy growth. Numerous scientific suggestions provided by academicians have been adopted by the government, such as the establishment of the National Natural Science Foundation of China in 1986, the setting up of 863 program in 1986, and the founding of CAE in 1994(Sheng and Han, 1999). Besides, to fully play the role as a scientific advisory body, CASAD strengthens the cooperation with government departments, universities and other research institutions. Some large-scale projects, like South—North Water Transfer Project, Qinghai—Tibet Railway Project, are all based on the evidence and scientific research from the joint efforts of academicians in CAS and CAE.

There are many channels for CAS academicians to support government decision-making. For instance, through research projects commissioned by government departments, research reports were provided initiatively to support the government, and ideas of its members could also be delivered to the top government leaders directly through special channels. To guarantee that CASAD could conduct its research projects independently, stable funding was allocated for its Members to focus on research areas where they are interested in. There are mainly two types of working patterns, one is to rely on the joint efforts of academicians from different divisions and disciplines to address more comprehensive issues, and the other is to assign individual division to deal with specific issues. The study process of CASAD[7] is demonstrated in Fig. 1.

To guarantee the high quality of the advisory work, the Consultation and Evaluation Committee and the Committee of each Academic Division invite experts from CAS and other external experts to review the report independently. After the review, the report[8] will be revised for several times. For those advisory work conducted initiatively by CAS, the advisory reports will be provided directly to related government departments. Also, the reports are sometimes delivered and released to the public.

One typical case of providing scientific advice initiatively is the formulation of "Science & Technology in China: A Roadmap to 2050", which was released in 2009. To address issues, such as energy, environmental and large population problems in China, CAS realized that further strategic priorities for science and technology development should be worked out with a long-term

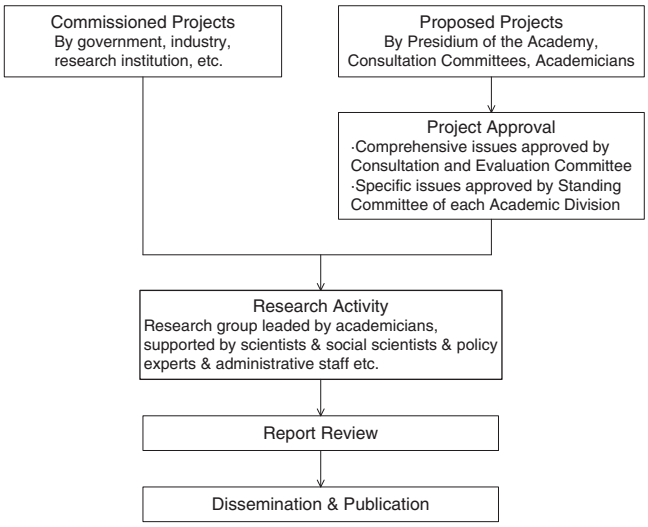


Figure 1 | Study process of CASAD.

view. In July 2007, the former CAS President Lu Yongxiang advocated CAS to lay out a science and technology roadmap for the next 20–30 years. According to the Executive Management Committee of CAS, the strategic research on science and technology roadmap for future development should be conducted to orchestrate the needs of both the nation and society, and target at three objectives: the growth of economy and national competiveness, the development of social harmony, and the sustainability between man and nature (Lu, 2010).

Relevant workshops were organized with the participation of scientists both within CAS and outside to discuss the research priorities and objectives. More than 300 experts, including 60 academicians and other scientific and technological experts, management experts and intelligence experts from around 80 research institutes of CAS, were involved in the process of drawing this roadmap. The final work includes a general report and 17 special ones on sub-fields such as energy, population health, mineral resources, space and ocean, information, materials, ecology and environment (CAS, 2009). The general report identifies crucial science and technology areas in the following 50 years in China, figures out the characteristics of these areas and objectives for their progress, and also sets out research agendas to achieve the progress. It provides sound evidence for government to make strategic decisions for development of science and technology in China. After the reports finished, they were delivered to both central government and local government and also were released to the public.

Apart from the above two modes presented in Fig. 1, there are special channels for academicians to present their ideas directly to the top national leaders. For instance, in May 1981, during the Fourth General Assembly of the CAS Members, 89 academicians jointly signed a letter, calling on the government to support basic research across the country. The suggestion of setting up the Science Foundation of CAS, which was later restructured as the National Natural Science Foundation of China, was seriously considered and adopted by the central government. During the late 1980s and the early 1990s, the tide of science and technology system reform waves across the nation, numerous scientific advisory reports on crucial issues were provided by academicians through this special channel to affect government's decisions on science and technology development, such as, the setting up of "National High Technology Research and Development Program (863 Program)" to follow the international advanced technology (Ru and Su, 2010), and some of them were completely adopted by central government.

Providing scientific advice to government mainly through the platform formed by research institutes. There are three major channels for research institutes to play the role as scientific advisory bodies. The first one is to jointly provide scientific advice on comprehensive issues. The second one is to conduct research works individually on specific issues. The third one is through the professional knowledge provided by prominent scholars as experts in scientific advisers group organized by government departments. In order to motivate research institutes to provide scientific advice to government, the number and impact of advisory reports is also used as an indicator to evaluate the institute's performance.[9]

The mode of providing scientific advice through research institutes can be seen from the draft of "National Major Function Oriented Zoning Plan". To remedy the situation that economy growth based on the excessive consumption of resources, the central government realized that it is necessary to develop a scientific and rational overall planning of regional development in China(Wu, 2015). Before National Development and Reform Commission (NDRC) commissioned CAS to conduct research into this issue, CAS had been working on it from 2003(Fan, 2013). The process that CAS plays its role as scientific advisory body to affect the government could be divided into three stages (Zhang, 2010).

Stage 1, expert groups were set up in CAS, which is led by three academicians and supported by research groups headed by Fan Jie from the Institute of Geographic Sciences and Natural Resources Research (IGSNRR), CAS. Fan Jie was also appointed as the Chief Scientist of the working group to fully in charge of the work.[10] After several communications with NDRC and scientific studies were conducted, the working group finally finished the scientific advisory report, dividing the national territory into five categories according to the area's level of economic development and resource & environmental bearing capability. The concept in the report was adopted by the government and accordingly the original five categories were replaced by "four main functional areas"(Zhang, 2010). The four areas are: prioritized development area, optimized development area, constrained development area and forbidden development area. Stage 2, based on the advisory report, NDRC commissioned the research group which was leaded by Fan Jie to work out how to develop the constrained development area. The method that put forward by the group was completely adopted in China's "11th Five-Year Development Guidelines (2006–2010)", which was guideline for China's next five years development of economy and society. Stage 3, the research group directly participated in formulating the "National Major Function Oriented Zoning Plan" as in charge of the first part to provide assessment criteria and methods of National Major Function Oriented Zoning.

Standardization and institutionalization of "science and technology think tanks"

Period III (2012 – present). Since the 18th National Congress of the Communist Party of China, the new government Xi Jinping has officially declared the concept of governance to push the reform of government management system, encouraging involvement of various non-government forces in governance. To improve the capability of consulting system, the central government extremely advocates the construction of "science and technology think tanks", identifying 25 high-quality think tanks[11] as pilot units in December, 2015. CAS is one of the 25 high-quality think tanks.

The central government states that it should decentralize its administrative authority and invite third-party organizations to evaluate the effect of its policies and measures. We assume that the third-party evaluation is one of the advisory activities, which is quite distinct due to its clear objectives and specific logic models. The third-party evaluation of government performance has been a critical measure taken by the central government to promote the reform of government management system (Xie, 2014), and the Premier of China's State Council, Li Keqiang claimed that this measure should be a regular mechanism to supervise the government to ensure that their polices and measures achieve the desired results.[12] Currently, the third-party evaluation reports are supposed to be presented to the State Council and also should be presented in the State Council executive meeting, which was hosted by Li Keqiang. Problems presented in the report should be solved by the heads of related government departments in a certain period of time.

In this context, the central government raised new expectations and requirements of CAS as the role of advisory body. During his visit to CAS in July 2013, Chinese President Xi Jinping urged the academy to take the lead in setting up a high-level national science and technology think tank. In response to the demands of the government, CAS has actively launched the government's third-party evaluation projects on the one hand, and promoted the standardization and institutionalization of its function of providing scientific advice to improve its capability on the other hand.

A number of projects were commissioned from CAS since September 2013 when the State Council started the third-party evaluation activities to supervise government performance. Some of them were commissioned by the State Council, such as the assessment of "Drinking Water Safety in Rural Area Program" and "National Poverty Reduction Program (jing zhun fu pin)" (Center, 2014; Wang and Li, 2015). Some were commissioned by ministries and commissions directly under the State Council, for instance, "Consultation and Deliberation of Whether China Should Participate in Square Kilometer Array (SKA) Program" was commissioned by Ministry of Science and Technology (MOST) and undertaken by the CASAD, "The Performance Assessment of China Agricultural Research System" project was commissioned by the Ministry of Finance and the Ministry of Agriculture to the Research Center of the Third-Party Evaluation of CAS.

Take the assessment of "Drinking Water Safety in Rural Area Program" as an example. In 2014, the State Council commissioned the IGSNRR of CAS to assess the performance of "Drinking Water Safety in Rural Area Program", which was conducted by the Ministry of Water Resources in the period of the "National Twelfth Five-year Plan(2011–2015)". CAS realized that it was a complicated project that required a large amount of human and scientific resources, and as such it authorized the CAS's Bureau of Development and Planning to be fully in charge of the project under the leadership of the president of CAS, Bai Chunli. On the basis of dozens of field researches, workshops and interviews with local farmers, experts from the IGSNRR and other CAS institutes assessed the performance of the program with the focus on the following aspects: the selection and protection of rural area's water sources, the management and maintenance of operation during and after project construction, water quality and quantity, the convenience and satisfaction of water-users and so on.

In contrast to the previous performance assessments of government department's programs, which were normally organized by the government itself with participation of related experts, this assessment project was conducted by an independent third party, CAS, which could perform with a relatively objective perspective to assess the performance of the government, to identify problems and put forward solutions through a large number of researches and investigations to support future implementation of the program.

To ensure that CAS could implement the third-party evaluation activities effectively and efficiently, efforts are made continually to explore the appropriate management systems and also to improve evaluation methods and approaches. Take "Consultation and Deliberation of Whether China Should Participate in SKA Program" for example, it was conducted by a special organization, the Research Center of the Third-Party Evaluation of CAS, rather than through the traditional consultation mechanism that working out the report after several communications and discussions with academicians, the new working pattern of "working group cooperating with expert group" was adopted in CAS to guarantee the high quality of advisory report.

The SKA is a large multi radio telescope project aimed to answer fundamental questions about the origin and evolution of the Universe, which is built in Australia and South Africa. It would have a total collecting area of approximately 1 km^2 and be able to survey the sky more than 10,000 times faster than ever before once built. However, it is a large-scale and comprehensive project which requires a large amount of money and scientific resources with the collaboration of multi-nations (https://en.wikipedia.org/wiki/Square_Kilometre_Array). It will require high performance and reliable research facilities, such as, high performance central computing engines and low cost broadband antenna, and these techniques will involve many research areas like the astronomy, radio, information science, mechanical engineering, computational mathematics and systems science and so on (http://www.most.gov.cn/mostinfo/xinxifenlei/fgzc/gfxwj/gfxwj2015/201512/t20151210_122801.htm).

In March 2015, MOST commissioned CASAD to take an assessment on whether China should participate in SKA program considering the ability and capacity of Chinese scientific resources. To guarantee the high quality of the advisory work, it was conducted by a special organization, the Research Center of the Third-Party Evaluation of CAS, one of the five supporting centers to support academicians to conduct advisory activities of the Institutes of Science and Development of CAS (CASISD), which was founded as an integration platform for advisory activities in CAS. During the advisory process, a new working mechanism was designed to maintain balance and objectivity throughout the work. After several consultations with related experts, an expert group, with 17 prestigious specialists from different perspectives and research areas, was set up, 15 of them are academicians from CAS and one from CAE. The research areas of the experts covered several subjects, for example, Astronomy, Physics, Electronics, Information, Earth Science, and also one expert is specialized in science and technology assessment. The leader of the expert group was chosen for his relatively neutral identity and full of experience.

To guarantee the high quality and the efficiency of the advisory work, a working group was also set up to support the scientific activities, such as collection of data and related materials, negotiation with experts, organization of workshops and seminars, drafting reports and other supporting works. The working group is comprised of staff from National Astronomical Observatories of CAS, National Science Library of CAS (NSLC), Bureau of Academic Divisions of CAS, Research Center of the Third-Party Evaluation of CAS and so forth. This new working pattern with the support of working group could avoid the shortcomings of the low quality caused by the lack of time of experts. The report was finished by the working group with the advisory and supporting from the experts, and was revised for

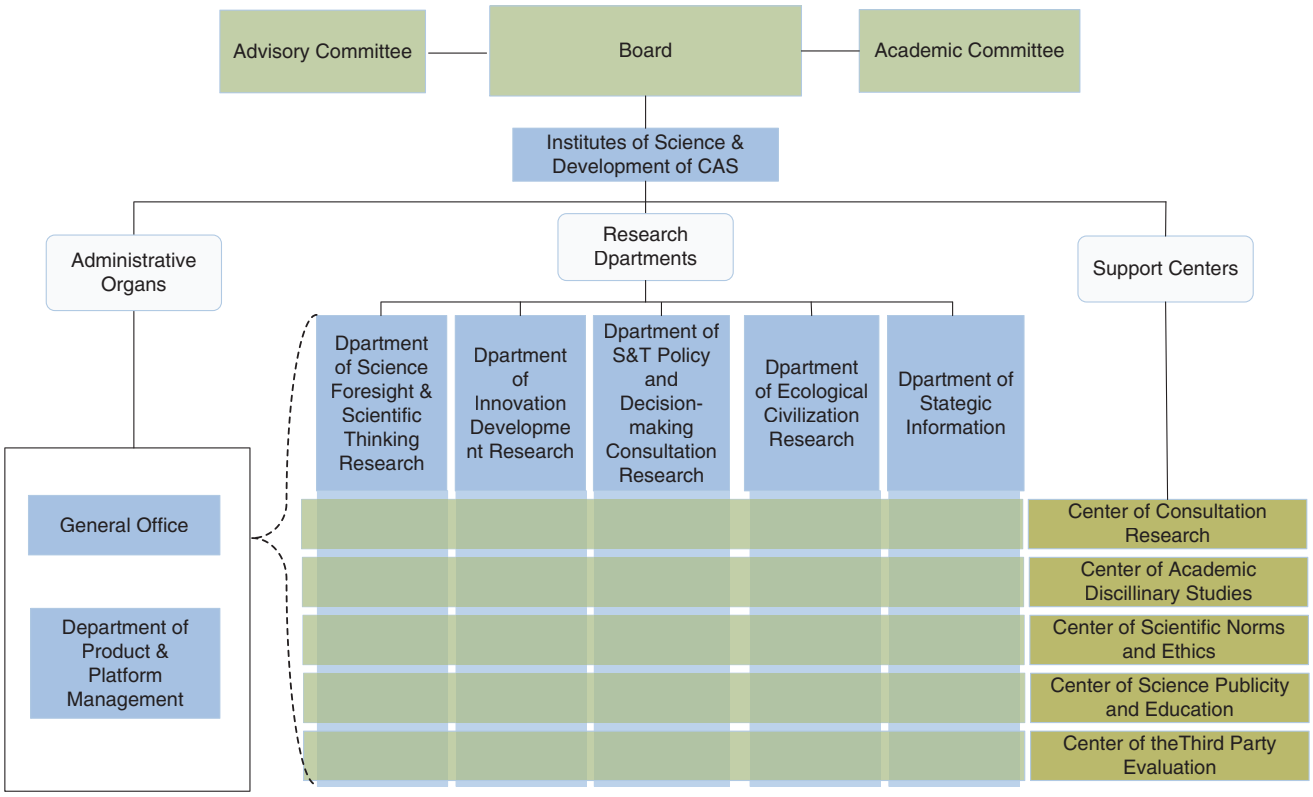


Figure 2 | Organizational structure of CASISD.
***Source*: internal report for 4th meeting of Presidency Office provided by Office of the CAS committee for Science Think-Tank Development.**

several times through communications with experts individually by email and face to face to reach a clear and balanced opinion. The final version of the report was presented to the government with the statement of opinions and evidence.

The new working mechanism of "expert group" cooperating with "working group" was designed by CAS to fully collect experts' opinions from different perspectives. This working pattern can make the best use of the time of experts and present the report in the language that can be understood by policy-makers, as the "working group" includes policy experts not just technology experts. Also, the final report presented clear opinions with sound evidence rather than just collections of different views, and that could be used directly by policymakers. This working mechanism of advisory is one of the new explorations of CAS to reform its advisory system.

As one of the critical measures taken by the central government to promote the reform of government management system, the third-party evaluation plays an important role to push the construction of a sound supervision mechanism of government, which requires the third party must be objective and scientific. From the perspective of objective, it does not mean that third party is absolutely disinterested organization, but to be objective with fully consideration of views from all the stakeholders. As the third party, CAS can play a relatively objective role in assessment of the performance of "Drinking Water Safety in Rural Area Program", whereas, it is one of the stakeholders in the "Consultation and Deliberation of Whether China Should Participate in SKA Program". But the latter program has been conducted successfully by CAS considering the balanced views of all related scientists from different perspectives to improve the objectivity and scientificity of the report.

In June 2015, based on the Institute of Policy and Management, Chinese Academy of Sciences, CASISD[13] was founded as a corporate body, with other research staff from IGSNRR, NSLC and the Institute for the History of Natural Sciences of CAS. CASISD is constructed as a research and supporting institute to support and assist academicians in providing strategic consultations for central government and also is constructed as a research and advisory platform to conduct research projects by integrating all the research resources of CAS. The relationship between CASISD and CASAD is similar as the relationship between NRC and National Academies, and Science Policy Center and the Royal Society. However, there are some differences, as CASISD could also integrate research forces from other research institutes across CAS. CASISD's organizational structure could be demonstrated in Fig. 2.

The chair of the board of CASISD is Bai Chunli, the president of CAS. There are two committees of CASISD, one is the Advisory Committee, which is made up of external experts and the other is the Academic Committee, the CAS Committee for Science Think-Tank Development, which is made up of internal experts of CAS. There are mainly five research departments of CASISD focusing on five different research areas based on the research institutes and research centers of CAS, orienting by their own research tasks and problems in their research frontier. The five departments are: Department of Science and Foresight and Scientific Thinking Research, Department of Innovation Development Research, Department of Science and Technology Policy and Decision-Making Consultation Research, Department of Ecological Civilization Research, and Department of Strategic Information. There are mainly five research centers to support academicians to carry out their advisory activities, disciplinary studies, scientific norms and ethics studies, science publicity, and third-party evaluation activities respectively. Apart from

supporting the academicians to carry out the third-party evaluations, the Research Center of the Third-Party Evaluation also support CAS to carry out research institutes' performance evaluation. CASISD also coordinates with University of CAS to push the integration of science and education. There is also a communication and data sharing platform to release the research productions of CASISD to the public.

As one of the 25 pilot units of the construction of high-quality think tanks in China, the CASISD is still on its way to explore and construct a sound mechanism in Chinese situation considering the experience of US NRC(National Research Council) and other advisory bodies in the world, and there are various issues, such as how to guarantee its independency and high quality of advisory report in China's current political and economic context, should be tackled before it reach the nation's expectation of being a high-caliber national science and technology think tank. There are many challenges for CAS to overcome compared with NRC and Royal Society, like how to guarantee independence of CAS with an extensive network of research institutes as "in house" resources.

Conclusion

We can draw some conclusions from what we have discussed in regard to the changing role of CAS in the Chinese scientific advisory system by means of an historical review. As the most prestigious academic research institute and a national academy, CAS has been playing a significant role in providing advice for government since its establishment. However, we divide it into three periods according to the different roles it played.

Period I (1949–1976): CAS was assigned as the top-level scientific advisor to Chinese government since its establishment, and its dual role of "scientific advisory body" and "science and technology administrative organ" especially evident through the process of making "The Twelve-Year National Long-term Outline for Science and Technology Development (1956–1967)".

Period II (1977–2011): after Cultural Revolution, scientific advice had never been in greater demand to tackle problems confronted in China's economic and social development and Chinese government also vigorously promoted the scientific and democratic construction of government's decision-making system, thus giving an extended period of scientific advice prosperity in CAS. During this period, a structured institutional system was formalized both in CASAD and research institutes of CAS and a large amount of high quality scientific advisory reports were produced.

Period III (2012 to present): with the promotion of reform in government management system, the new government extremely advocates the construction of "science and technology think tanks", stating that the central government should decentralize its administrative authority and invite third-party organization to evaluate government performance. Under this background, the construction of think tank ushered in a new era of standardized and institutionalized system. CAS, as the third-party organization, actively explores a new mode to give full play to its unique advantages, integrating research resources from the Academic Divisions and Research Institutes, to further develop its consulting approaches and improve its consulting capability and organizational structure.

Considering the prominence of CAS in Chinese scientific advisory system, the changing role of CAS reflects, to some extent, the characteristics of China's scientific advice and its development. Two points could be drawn out.

The first point is that the Chinese government increasingly pays more attention to the construction of science and technology think tanks and the structure of the scientific advisory system has greatly improved. Since the 16th National Congress of the Communist Party of China, the government has been repeatedly emphasizing that policy-making mechanisms should be reformed and improved so as to ensure policy-making is more scientific and democratic. After the 1980s, various advisory bodies sprang up to conduct research and explorations on technological, industrial and major strategic policies so as to support government policy-making (Li *et al.*, 2015). In recent years, with the advocating of the concept of governance to push the reform of government management system, the central government attached great importance on the third-party organization to push the construction of government's supervision mechanism. However, there are still challenges to keep the balance of objectivity and influence of think tanks in China, and the influence power of non-governmental organization on government decision-making is still much weaker than those traditional think tanks. Moreover, there is much more to do to improve the objectivity and independence of the third-party organization in China. It is necessary for the central government and the politicians to understand the importance and respect the independence of third-party organizations to play their due role. Also, third-party organizations, such as research institutions, universities, associations and societies, and NGOs, should strictly control the process of their study to maintain the high quality of their research reports.

The second point is that even though the Chinese government has been promoting the development of scientificity of decision making since 1980s, there are still numerous problems to be addressed, like the lack of laws to guarantee the independence of think tanks. Other problems like how to adjust organizational structure and improve evaluation approach to work efficiently in Chinese situation, how to guide and motivate scientists and engineers and other experts to engage in scientific advice activity, are still should be tackled for think tanks to improve the quality of their scientific advisory activities. Currently most CAS scientists prefer to do more research experiments and publish more research papers, rather than involving themselves in advisory activities and providing advisory reports. It is encouraging that CAS is exploring solutions to all these problems and advancing itself to the national high-level science and technology think tank as required by Chinese President Xi Jinping.

Notes

1 Brief introduction to CAS, http://english.cas.cn/about_us/introduction/201501/t20150114_135284.shtml.
2 There was no MOST (Ministry of Science and Technology of the People's Republic of China) or other regulatory authority on science and technology in China at that time, so CAS was performed as an administrative organ to be in charge of the national development of science and technology and the management of scientific resources. However, the role of CAS as an administrative organ was not clear at that time.
3 The State Planning Committee of Science was founded to take full charge of the formulating of the Twelve-Year Plan. However, given its small group members, individual area plan designing was undertaken by individual groups in different areas. For example, the State Technology Commission was in charge of the plan designing in industries and transportations, CAS was in charge of the development of basic sciences, Department of Agriculture and Department of Health was in charge of plan designing in agricultures and medical sciences respectively.
4 The State Technology Commission was founded in 1956 to fully in charge of working out the 5 years and long-term plan for technology development; The State Planning Committee of Science was founded later in the same year to be in charge of working out the "The Twelve-Year National Long-term Outline for Science and Technology Development (1956–1967)" ; then in 1958, the two commissions were merged into one, named the State Science and Technology Commission, which was renamed as MOST in1998 after the reform of the scientific management system in China.
5 The activities of the Academic Divisions came to a complete halt for more than 10 years.
6 Review of the history of General Assembly of the members of the CAS, http://www.cas.cn/zt/hyzt/zgkxydswcysdh/ljysdhhg/.

7 Regulations on the Management of Consulting Projects of the Chinese Academy of Sciences, http://www.casad.ac.cn/doc/14870.html.
8 CAS generally provides recommendations on policies and sometimes it also offers policy options.
9 Advisory report was one of the CAS institute performance indicators at that time. Other indicators include publications, patents, projects, talents and rewards and so on.
10 Zhang Fuxing. CAS provided scientific advice for the formulation of "National Main Functional Area Plan". *China Science Daily*, 2010-7-6.http://www.cas.cn/xw/cmsm/201007/t20100706_2893082.shtml.
11 25 high-quality think tanks as the pilot units, http://www.js.xinhuanet.com/2015-12/04/c_1117358781.htm[2016-3-20].
12 Wang ziyue. Reports of third-party evaluation were presented to State Council: it was declared that the third-party evaluation of government performance should be a regular mechanism held every year. *Yicai Daily*, http://finance.sina.com.cn/china/20140901/021720173303.shtml[2014-9-1].
13 Brief introduction of CASISD, http://english.casisd.cn/about/introduction/index.html.

References

Arimoto T and Sato Y (2012) Rebuilding public trust in science for policy-making. *Science*; **337** (6099): 1176–1177.
Australian & Government Department of Industry, I., Science, Research and Tertiary Education. (2012) APS200 Project: The Place of Science in Policy Development in the Public Service. Canberra.
Blair P (2006) Scientific advice for policy in the United States: Lessons from the National Academies and the former Congressional Office of Technology Assessment. Berlin, Germany.
Cao C (2004) *China's Scientific Elite*. Routledge: London.
Cao C and Suttmeier RP (1999) China's "brain bank": Leadership and Elitism in Chinese science and engineering. *Asian Survey*; **39** (3): 525–559.
CAS, B. O. B. S. (2009) Excerpts of the report of "science & technology in China: A roadmap to 2050. *Frontier Science*; **3**(11): 4–19.
Center, I. P. O. M. (2014) Using third-party evaluation approach to supervise state council for the first time. *Information for Deciders Magazine*; (35): 26–27.
Cheng S (2005) Vigorously promoting the development of soft science to advance the scientific and democratic construction of decision making. *China Soft Science*; (4): 1–6.
Collins P (2011) Quality control in scientific policy advice: The experience of the royal society. In: Lentsch J and Weingart P (eds). *The Politics of Scientific Advice: Institutional Design for Quality Assurance*. Cambridge University Press: Cambridge, UK, pp 334–341.
Doubleday R and Wilsdon J (eds) (2013) *Future directions for scientific advice in Whitehall*. Alliance for Useful Evidence & Cambridge Centre for Science and Policy: London.
Fan J (2013) The strategy of major function oriented zoning and the optimization of territorial development patterns. *Bulletin of the Chinese Academy of Sciences*; **28**(2): 54–67.
Glynn S, Cunningham P and Flanagan K (2002) Typifying Scientific Advisory Structures and Scientific Advice Production Methodologies (TSAS). Draft Final Report, PREST, University of Manchester: UK.
Hou J (2009) Development history and its future improvement approaches of CPC's scientific and democratic decision-making system. *Weishi*; (3): 23–25.
Hu W (2006) The 12-year long-term science and technology development plan: Planning, effect and some enlightenments. *Bulletin of Chinese Academy of Sciences*; (3): 207–212.
Jasanoff S (2009) *The Fifth Branch: Science Advisers as Policymakers*. Harvard University Press: Cambridge, MA.
Li A (1999) A milestone in the history of New China's science and technology development—The Twelve-Year National Long-term Outline for Science and Technology Development (1956–1967). *Science News*; 30(28): .
Li H (1991) Historical review of "the twelve-year national long-term outline for science and technology development (1956–1967)". *Truth Seeking*; (4): 20–21.
Li X, Yang K, Zhang X, Liu X and Liu Z (2015) *A Comparative Study on How Todevelop an Evidencebased Approach to Policy-Making*. Science Press: Beijing, China.
Lu Y (Editor in Chief) (2010) *Science & Technology in China: A roadmap to 2050*; Strategic General Report of the Chinese Academy of Sciences. Science Press: Beijing, China.
OECD. (2015) Scientific Advice for Policy Making: The Role and Responsibility of Expert Bodies and Individual Scientists. OECD Science, Technology and Industry Policy Papers, No. 21, OECD Publishing: Paris, http://dx.doi.org/10.1787/5js33l1jcpwb-en.
Qian B (2010) Formation of science and technology management system in the early years of new China. *Contemporary China History Studies*; 17(3): 44-51 +125-126.
Ru P and Su J (2010) Research on the influence of scientists in science and technology decision making—Taking 863 program as example. *China Soft Science*; (10): 86–92.
Sheng H and Han C (1999) Academic Divisions of Chinese Academy of Sciences: National Science Think Tank, National Highest Scientific Advisory Body. *Science News*; (32): 22–23.
Wang J (2008) Research review of ten years history of government's scientific and democratic decision making system—From the 15th national congress to the 17th national congress. *Journal of Xinyang Normal University (Philosophy and Social Science Edition)*; (4): 21–24.
Wang J and Li Y (2015) The third-party evaluation report of supervision of state council's policies and measures conducted by geographers. *Acta Geographica Sinica*; (10): 1694–1695.
Wang Y (2015) Reflections on the history of the CASAD at its 60 anniversary. *Bulletin of Chinese Academy of Sciences*; **30** (3): 414–420.
Wei L (2006) Research on the evolution of function and mission of Academic Divisions of CAS. *Management and Review of Social Sciences*; (1): 35–46.
Wilsdon J, Allen K and Paulavets K (2014) Science Advice to Governments. A briefing paper for the Auckland conference, 28–29 August. Auckland.
Wu F (2015) *Planning for Growth: Urban and Regional Planning in China*. Routledge: New York, pp 119–142.
Wu H (1994) *The Fifty Years History of Science and Technology in China from 1940 to 1990 (Ke ji zhan xian wu shi nian)*. Scientific and Technical Documentation Press: Beijing, China, pp 159–167.
Xie W (2014) Government decentralization scoring by third-party evaluation report. *China Economic Weekly*; (43): 62–65.
Yang W and Zhang M (2007) Making and implementation of the first plan of scientific and technological development in new China and the histrocial experience. *CPC History Studies*; (6): 42–49.
Zhang F (2010) CAS provided scientific advice for the formulation of "National Main Functional Area Plan". *China Science Daily*. http://www.cas.cn/xw/cmsm/201007/t20100706_2893082.shtml
Zheng C, Yang D and Jingru L (2013) An analysis of the model of scien-tech organisations' participation in decision making and consultancy. *First Resource*; (1): 131–147.

Data availability

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Additional information

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available at http://www.palgrave-journals.com/pal/authors/rights_and_permissions.html

How to cite this article: Li X, Yang K and Xiao X (2016) Scientific advice in China: the changing role of the Chinese Academy of Sciences. *Palgrave Communications*. 2:16045 doi: 10.1057/palcomms.2016.45.

Exhibit

37

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Virology®

VIRUS-CELL INTERACTIONS

Molecular Mechanism for Antibody-Dependent Enhancement of Coronavirus Entry

Yushun Wan,[a] Jian Shang,[a] Shihui Sun,[b] Wanbo Tai,[c] Jing Chen,[d] Qibin Geng,[a] Lei He,[b] Yuehong Chen,[b] Jianming Wu,[a] Zhengli Shi,[d] Yusen Zhou,[b] Lanying Du,[c] Fang Li[a]

[a]Department of Veterinary and Biomedical Sciences, College of Veterinary Medicine, University of Minnesota, Saint Paul, Minnesota, USA
[b]Laboratory of Infection and Immunity, Beijing Institute of Microbiology and Epidemiology, Beijing, China
[c]Lindsley F. Kimball Research Institute, New York Blood Center, New York, New York, USA
[d]Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, Hubei Province, China

Yushun Wan and Jian Shang contributed equally to this work. Author order was determined by the time of joining the project.

ABSTRACT Antibody-dependent enhancement (ADE) of viral entry has been a major concern for epidemiology, vaccine development, and antibody-based drug therapy. However, the molecular mechanism behind ADE is still elusive. Coronavirus spike protein mediates viral entry into cells by first binding to a receptor on the host cell surface and then fusing viral and host membranes. In this study, we investigated how a neutralizing monoclonal antibody (MAb), which targets the receptor-binding domain (RBD) of Middle East respiratory syndrome (MERS) coronavirus spike, mediates viral entry using pseudovirus entry and biochemical assays. Our results showed that MAb binds to the virus surface spike, allowing it to undergo conformational changes and become prone to proteolytic activation. Meanwhile, MAb binds to cell surface IgG Fc receptor, guiding viral entry through canonical viral-receptor-dependent pathways. Our data suggest that the antibody/Fc-receptor complex functionally mimics viral receptor in mediating viral entry. Moreover, we characterized MAb dosages in viral-receptor-dependent, Fc-receptor-dependent, and both-receptors-dependent viral entry pathways, delineating guidelines on MAb usages in treating viral infections. Our study reveals a novel molecular mechanism for antibody-enhanced viral entry and can guide future vaccination and antiviral strategies.

IMPORTANCE Antibody-dependent enhancement (ADE) of viral entry has been observed for many viruses. It was shown that antibodies target one serotype of viruses but only subneutralize another, leading to ADE of the latter viruses. Here we identify a novel mechanism for ADE: a neutralizing antibody binds to the surface spike protein of coronaviruses like a viral receptor, triggers a conformational change of the spike, and mediates viral entry into IgG Fc receptor-expressing cells through canonical viral-receptor-dependent pathways. We further evaluated how antibody dosages impacted viral entry into cells expressing viral receptor, Fc receptor, or both receptors. This study reveals complex roles of antibodies in viral entry and can guide future vaccine design and antibody-based drug therapy.

KEYWORDS antibody-dependent enhancement of viral entry, MERS coronavirus, SARS coronavirus, spike protein, neutralizing antibody, viral receptor, IgG Fc receptor, antibody-dependent enhancement of viral entry

Antibody-dependent enhancement (ADE) occurs when antibodies facilitate viral entry into host cells and enhance viral infection in these cells (1, 2). ADE has been observed for a variety of viruses, most notably flaviviruses (e.g., dengue virus) (3–6). It has been shown that when patients are infected by one serotype of dengue virus (i.e.,

Citation Wan Y, Shang J, Sun S, Tai W, Chen J, Geng Q, He L, Chen Y, Wu J, Shi Z, Zhou Y, Du L, Li F. 2020. Molecular mechanism for antibody-dependent enhancement of coronavirus entry. J Virol 94:e02015-19. https://doi.org/10.1128/JVI.02015-19.

Editor Tom Gallagher, Loyola University Chicago

Address correspondence to Lanying Du, LDu@nybc.org, or Fang Li, lifang@umn.edu.

Received 27 November 2019
Accepted 4 December 2019
Accepted manuscript posted online 11 December 2019
Published 14 February 2020

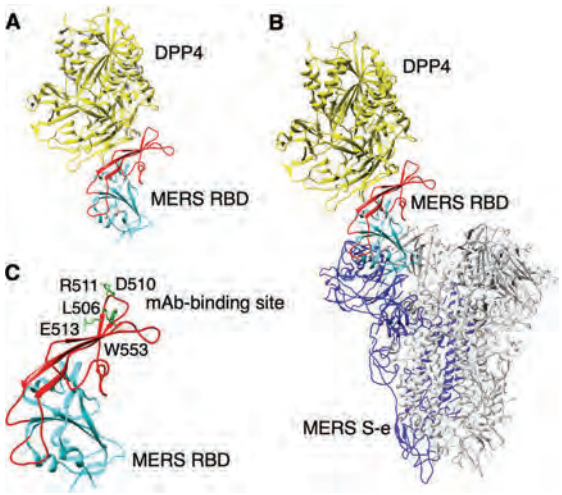


FIG 1 Structural similarity between DPP4 and MAb in binding MERS-CoV spike. (A) Tertiary structure of MERS-CoV RBD in complex with DPP4 (PDB code 4KR0) (30). DPP4 is colored yellow. RBD is colored cyan (core structure) and red (receptor-binding motif). DPP4 binds to the receptor-binding motif of the RBD. (B) Modeled structure of MERS-CoV S-e in complex with DPP4. S-e is a trimer (PDB code 5X5F): one monomeric subunit, whose RBD is in the standing-up conformation, is colored blue, and the other two monomeric subunits, whose RBDs are in the lying-down conformation, are colored gray (18). To generate the structural model of the S-e in complex with DPP4, the RBD in panel A was structurally aligned with the standing-up RBD in the S-e trimer. (C) Tertiary structure of MERS-CoV RBD (PDB code 4L3N) (64). Critical MAb-binding residues were identified through mutagenesis studies (48) and are shown as green sticks.

primary infection), they produce neutralizing antibodies targeting the same serotype of the virus. However, if they are later infected by another serotype of dengue virus (i.e., secondary infection), the preexisting antibodies cannot fully neutralize the virus. Instead, the antibodies first bind to the virus and then bind to the IgG Fc receptors on immune cells and mediate viral entry into these cells. A similar mechanism has been observed for HIV and Ebola viruses (7–10). Thus, subneutralizing antibodies (or non-neutralizing antibodies in some cases) are responsible for ADE of these viruses. Given the critical roles of antibodies in host immunity, ADE causes serious concerns in epidemiology, vaccine design, and antibody-based drug therapy. This study reveals a novel mechanism for ADE in which fully neutralizing antibodies mimic the function of viral receptor in mediating viral entry into Fc receptor-expressing cells.

Coronaviruses are a family of large, positive-stranded, and enveloped RNA viruses (11, 12). Two highly pathogenic coronaviruses, severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV), cause lethal infections in humans (13–16). An envelope-anchored spike protein guides coronavirus entry into host cells (17). As a homotrimer, the spike contains three receptor-binding S1 subunits and a trimeric membrane fusion S2 stalk (18–25). This state of the spike on the mature virions is called "prefusion." SARS-CoV and MERS-CoV recognize angiotensin-converting enzyme 2 (ACE2) and dipeptidyl peptidase 4 (DPP4), respectively, as their viral receptors (26–28). Their S1 subunits each contain a receptor-binding domain (RBD) that mediates receptor recognition (29, 30) (Fig. 1A). The RBD is located on the tip of the spike trimer and is present in two different states: standing up and lying down (18, 21) (Fig. 1B). Binding to a viral receptor can stabilize the RBD in the standing-up state (20). Receptor binding also triggers the spike to undergo further conformational changes, allowing host proteases to cleave at two sites sequentially: first at the S1/S2 boundary (i.e., S1/S2 site) and then within S2 (i.e., S2′ site) (31, 32). Proteolysis of the spike can take place during viral maturation (by proprotein convertases), after viral release (by extracellular proteases), after viral attachment (by cell surface proteases), or after viral endocytosis (by lysosomal proteases) (33–39). After two protease cleavages, S1 dissociates and S2 undergoes a dramatic structural change to

fuse host and viral membranes; this membrane fusion state of the spike is called "postfusion" (40, 41). Due to the recent progress toward understanding the receptor recognition and membrane fusion mechanisms of coronavirus spikes, coronaviruses represent an excellent model system for investigating ADE of viral entry.

ADE has been observed for coronaviruses. Several studies have shown that sera induced by SARS-CoV spike enhance viral entry into Fc receptor-expressing cells (42–44). Further, one study demonstrated that unlike receptor-dependent viral entry, serum-dependent SARS-CoV entry does not go through the endosome pathway (44). Additionally, it has long been known that immunization of cats with feline coronavirus spike leads to worsened future infection due to the induction of infection-enhancing antibodies (45–47). However, detailed molecular mechanisms for ADE of coronavirus entry are still unknown. We previously discovered a monoclonal antibody (MAb) (named Mersmab1) which has strong binding affinity for MERS-CoV RBD and efficiently neutralizes MERS-CoV entry by outcompeting DPP4 (48); this discovery allowed us to comparatively study the molecular mechanisms for antibody-dependent and receptor-dependent viral entries.

In this study, we examined how Mersmab1 binds to MERS-CoV spike, triggers the spike to undergo conformational changes, and mediates viral entry into Fc receptor-expressing cells. We also investigated the pathways and antibody dosages for Mersmab1-dependent and DPP4-dependent viral entries. Our study sheds lights on the mechanisms of ADE and provides insight into vaccine design and antibody-based antiviral drug therapy.

RESULTS

Antibody-dependent enhancement of coronavirus entry. To investigate ADE of coronavirus entry, we first characterized the interactions between Mersmab1 (which is a MERS-CoV RBD-specific MAb) and MERS-CoV spike using biochemical methods. First, enzyme-linked immunosorbent assay (ELISA) was performed between Mermab1 and MERS-CoV RBD and between Mersmab1 and MERS-CoV spike ectodomain (S-e) (Fig. 2A). To this end, Mersmab1 (which was in excess) was used to coat the ELISA plate, and gradient amounts of recombinant RBD or S-e were added for detection of potential binding to Mersmab1. The results showed that both the RBD and S-e bound to Mersmab1. S-e bound to Mersmab1 more tightly than the RBD did, likely due to the multivalent effects associated with the trimeric state of S-e. Second, we prepared Fab from Mersmab1 using papain digestion and examined the binding between Fab and S-e using ELISA. Recombinant S-e (which was in excess) was used to coat the ELISA plate, and gradient amounts of Fab or Mersmab1 were added for detection of potential binding to S-e. The results showed that both Fab and Mersmab1 bound to S-e (Fig. 2B). Mersmab1 bound to S-e more tightly than Fab did, also likely due to the multivalent effects associated with the dimeric state of Mersmab1. Third, a flow cytometry assay was carried out to detect the binding between S-e and the DPP4 receptor and among S-e, Mersmab1, and CD32A (which is an Fc receptor). To this end, DPP4 or CD32A was expressed on the surface of human HEK293T cells (human kidney cells), and recombinant S-e was added for detection of potential binding to one of the two receptors in the absence or presence of Mersmab1. The results showed that without Mersmab1, S-e bound to DPP4 only; in the presence of Mersmab1, S-e bound to CD32A (Fig. 2C). As a negative control, a SARS-CoV RBD-specific MAb (49) did not mediate the binding of S-e to CD32A. The cell surface expressions of both DPP4 and CD32A were measured and used for calibrating the flow cytometry result (Fig. 2D), demonstrating that the direct binding of S-e to DPP4 is stronger than the indirect binding of S-e to CD32A through Mersmab1. Overall, these biochemical results reveal that Mersmab1 not only directly binds to the RBD region of MERS-CoV S-e but also mediates the indirect binding interactions between MERS-CoV S-e and the Fc receptor.

Next, we investigated whether Mersmab1 mediates MERS-CoV entry into Fc receptor-expressing cells. To this end, we performed a MERS-CoV pseudovirus entry assay, in which retroviruses pseudotyped with MERS-CoV spike (i.e., MERS-CoV pseu-

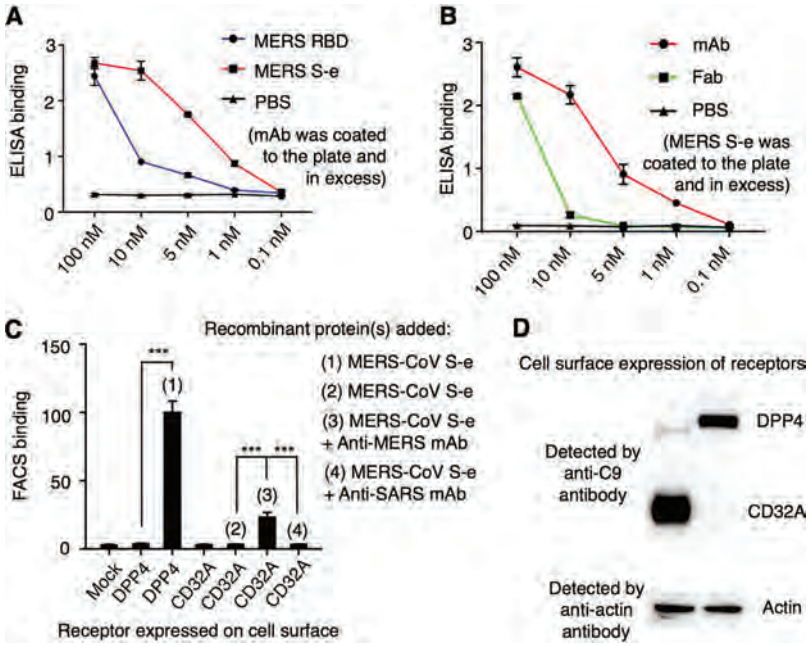


FIG 2 Interactions between coronavirus spike and RBD-specific MAb. (A) ELISA for detection of the binding between MERS-CoV RBD-specific MAb (i.e., Mersmab1) and MERS-CoV spike ectodomain (S-e). Mersmab1 was precoated on the plate, and recombinant S-e or RBD was added subsequently for ELISA. Binding affinities were characterized as ELISA signal at an optical density (OD) at 450 nm. PBS was used as a negative control. (B) ELISA for detection of the binding between Fab of Mersmab1 and MERS-CoV S-e. Recombinant S-e was used to precoat the plate, and Mersmab1 or Fab was added subsequently for ELISA. (C) Flow cytometry for detection of the binding between MERS-CoV S-e and DPP4 receptor and among S-e, Mersmab1, and CD32A (i.e., Fc receptor). Cells expressing DPP4 or CD32A were incubated with S-e alone, S-e plus Mersmab1, or S-e plus a SARS-CoV RBD-specific MAb (i.e., 33G4). Fluorescence-labeled anti-His_6 antibody was added to target the C-terminal His_6 tag on S-e. Cells were analyzed using fluorescence-activated cell sorting (FACS). (D) The expression levels of cell-membrane-associated DPP4 and CD32A were characterized using Western blotting targeting their C-terminal C9 tag and then used to normalize the binding affinity as measured in panel C. As an internal control, the expression level of cellular actin was measured using an anti-actin antibody. All of the experiments were repeated at least three times, with similar results, and representative results are shown. Error bars indicate SD ($n = 5$). Statistical analyses were performed as a one-tailed *t* test. ***, $P < 0.001$. Mersmab1 and its Fab both bind to MERS-CoV RBD and S-e.

doviruses) were used to enter human cells expressing CD32A on their surface. The main advantage of pseudovirus entry assay is to focus on the viral entry step (which is mediated by MERS-CoV spike) by separating viral entry from the other steps of viral infection cycles (e.g., replication, packaging, and release). We tested three different types of Fc receptors: CD16A, CD32A, and CD64A; each of these Fc receptors was exogenously expressed in HEK293T cells. We also tested macrophages in which mixtures of Fc receptors were endogenously expressed. The absence of Mersmab1 served as a control for Mersmab1 (a nonneutralizing MAb would be appropriate as another control for Mersmab1, but we do not have access to any nonneutralizing MAb). The results showed that in the absence of Mersmab1, MERS-CoV pseudoviruses could not enter Fc receptor-expressing cells; in the presence of Mersmab1, MERS-CoV pseudoviruses demonstrated significant efficiency in entering CD32A-expressing HEK293T cells and macrophages (Fig. 3A). In comparison, in the absence of Mersmab1, MERS-CoV pseudoviruses entered DPP4-expressing HEK293T cells efficiently, but the entry was blocked effectively by Mersmab1 (Fig. 3A). In control experiments, anti-SARS MAb did not mediate MERS-CoV pseudovirus entry into Fc receptor-expressing HEK293T cells or macrophages, nor did it block MERS-CoV pseudovirus entry into DPP4 receptor-expressing HEK293T cells (Fig. 3A). In another set of control experiments, we showed that neither the Fc nor the Fab portion of Mersmab1 could mediate MERS-CoV pseudovirus entry into Fc receptor-expressing HEK293T cells or macrophages (Fig. 3B),

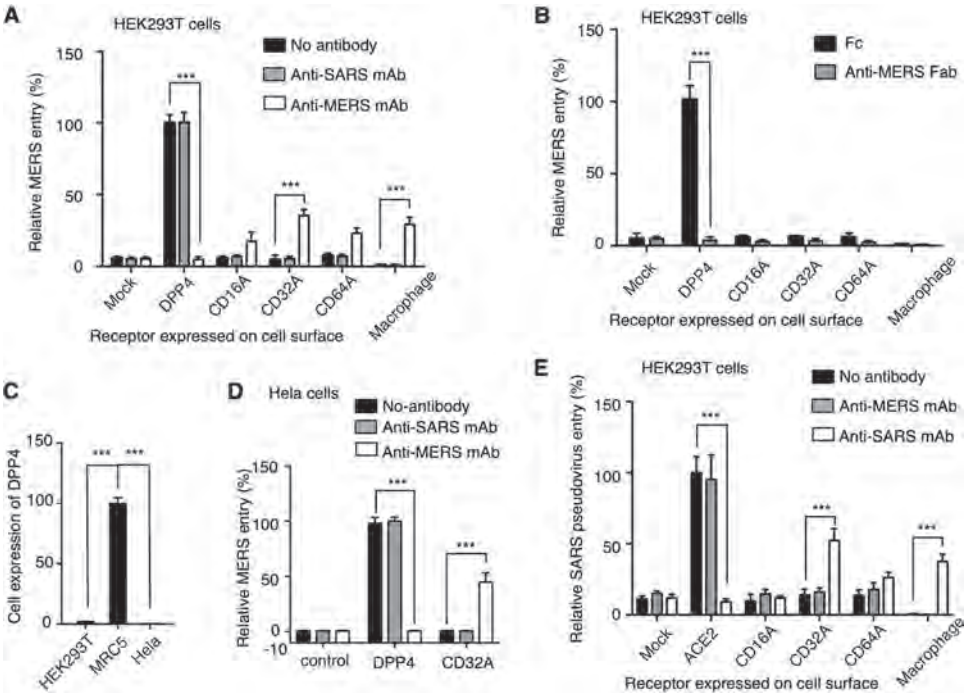


FIG 3 Antibody-dependent enhancement of coronavirus entry. (A) Antibody-mediated MERS-CoV pseudovirus entry into human cells. The human cells included HEK293T cells exogenously expressing DPP4, HEK293T cells exogenously expressing one of the Fc receptors (CD16A, CD32A, or CD64A), and macrophages (induced from THP-1 monocytes) endogenously expressing a mixture of Fc receptors. The antibody was Mersmab1. An anti-SARS MAb (i.e., 33G4) was used as a negative control. Efficiency of pseudovirus entry was characterized by luciferase activities accompanying entry. HEK293T cells not expressing any viral receptor or Fc receptor were used as a control. (B) Fc- or Fab-mediated MERS-CoV pseudovirus entry into human cells. The Fc or the Fab portion of Mersmab1 was used in MERS-CoV pseudovirus entry performed as for panel A. (C) Expression levels of DPP4 receptor in different cell lines. Total RNA was extracted from three different cell lines: HEK293T, MRC5, and HeLa. Then qRT-PCR was performed on the total RNAs from each cell line. The expression level of DPP4 in each cell line is defined as the ratio between the RNA of DPP4 and the RNA of glyceraldehyde-3-phosphate dehydrogenase (GAPDH). (D) Antibody-mediated MERS-CoV pseudovirus entry into HeLa cells that do not express DPP4 receptor. The experiments were performed in the same way as for panel A, except that HeLa cells replaced HEK293T cells. (E) Antibody-mediated SARS-CoV pseudovirus entry into human cells. DPP4 and Mersmab1 were replaced by ACE2 and 33G4, respectively. Mersmab1 was used as a negative control. All of the experiments were repeated at least three times, with similar results, and representative results are shown here. Error bars indicate SD ($n = 4$). Statistical analyses were performed as a one-tailed t test. ***, $P < 0.001$. RBD-specific MAbs mediate ADE of coronavirus entry while blocking viral-receptor-dependent coronavirus entry.

suggesting that both the Fc and Fab portions of anti-MERS MAb are required for antibody-mediated viral entry. The above-mentioned DPP4-expressing HEK293T cells were induced to exogenously express high levels of DPP4. To detect background expression levels of DPP4, we performed reverse transcription-quantitative PCR (qRT-PCR) on HEK293T cells. The result showed that HEK293T cells express very low levels of DPP4 (Fig. 3C). In comparison, MRC5 cells (human lung cells) express high levels of DPP4, whereas HeLa cells (human cervical cells) do not express DPP4 (Fig. 3C). Because of the comprehensive control experiments that we performed, the very low endogenous expression of DPP4 in HEK293T cells should not affect our conclusions. Nevertheless, we confirmed the above results using HeLa cells that do not express DPP4 (Fig. 3D). Overall, our results reveal that Mersmab1 mediates MERS-CoV entry into Fc receptor-expressing cells but blocks MERS-CoV entry into DPP4-expressing cells.

To expand the above-described observations to another coronavirus, we investigated ADE of SARS-CoV entry. We previously identified a SARS-CoV RBD-specific MAb, named 33G4, which binds to the ACE2-binding region of SARS-CoV RBD (49, 50); this MAb was examined for its potential capability to mediate ADE of SARS-CoV entry (Fig. 3E). The result showed that 33G4 mediated SARS-CoV pseudovirus entry into CD32A-expressing cells but blocked SARS-CoV pseudovirus entry into ACE2-expressing cells.

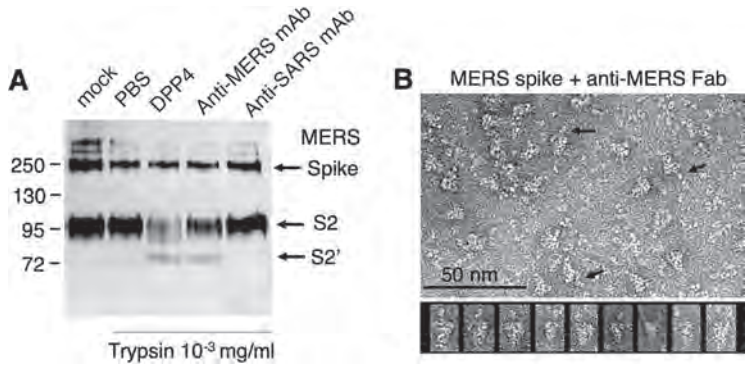


FIG 4 Antibody-induced conformational changes of coronavirus spike. (A) Purified MERS-CoV pseudoviruses were incubated with recombinant DPP4, MAb, or PBS and then treated with trypsin. Samples were subjected to Western blotting. MERS-CoV spike and its cleaved fragments (all of which contained a C-terminal C9 tag) were detected using an anti-C9 tag monoclonal antibody. Both DPP4 and Mersmab1 triggered conformational changes of MERS-CoV spike, allowing it to cleaved at the S2′ site by trypsin. (B) Negative-stain electron microscopic analysis of MERS-CoV S-e in complex with the Fab of Mersmab1. Both a field of particles and windows of individual particles are shown. Black arrows indicate S-e-bound Fabs. According to previous studies (18, 20, 21), the Fab-binding site on the trimeric S-e is accessible only when the RBD is in the standing-up position.

Therefore, both the MERS-CoV RBD-specific MAb and the SARS-CoV RBD-specific MAb can mediate the entry of the respective coronavirus into Fc receptor-expressing human cells while blocking the entry of the respective coronavirus into viral-receptor-expressing human cells. For the remainder of this study, we selected the MERS-CoV RBD-specific MAb Mersmab1 for in-depth analysis of ADE.

Molecular mechanism for antibody-dependent enhancement of coronavirus entry. To understand the molecular mechanism of ADE, we investigated whether Mersmab1 triggers any conformational change of MERS-CoV spike. It was shown previously that DPP4 binds to MERS-CoV spike and stabilizes the RBD in the standing-up position (Fig. 1A and B), resulting in a weakened spike structure and allowing the S2′ site to become exposed to proteases (51). We repeated this experiment: MERS-CoV pseudoviruses were incubated with DPP4 and then subjected to trypsin cleavage (Fig. 4A). The results showed that during the viral packaging process, virus surface-anchored MERS-CoV spike molecules were cleaved at the S1/S2 site by proprotein convertases; in the absence of DPP4, the spike molecules could not be cleaved further at the S2′ site by trypsin. These data suggest that only the S1/S2 site, and not the S2′ site, was accessible to proteases in the free form of the spike trimer. In the presence of DPP4, a significant amount of MERS-CoV spike molecules were cleaved at the S2′ site by trypsin, indicating that DPP4 binding triggered a conformational change of MERS-CoV spike to expose the S2′ site. Interestingly, we found that Mersmab1 binding also allowed MERS-CoV spike to be cleaved at the S2′ site by trypsin. As a negative control, the SARS-CoV RBD-specific MAb did not trigger MERS-CoV spike to be cleaved at the S2′ site by trypsin. Hence, like DPP4, Mersmab1 triggers a conformational change of MERS-CoV spike to expose the S2′ site for proteolysis.

We further analyzed the binding between Mersmab1 and MERS-CoV S-e using negative-stain electron microscopy (EM). We previously demonstrated through mutagenesis studies that Mersmab1 binds to the same receptor-binding region on MERS-CoV RBD as DPP4 does (Fig. 1C) (48). Because full-length Mersmab1 (which is a dimer) triggered aggregation of S-e (which is a trimer), we prepared the Fab part (which is a monomer) of Mersmab1, detected the binding between Fab and S-e (Fig. 2B), and used Fab in the negative-stain EM study. The results showed that Fab bound to the tip of the S-e trimer, where the RBD is located (Fig. 4B). Due to the limited resolution of negative-stain EM, we could not clearly see the conformation of the Fab-bound RBD. However, based on previous studies, the receptor-binding site on the RBD in the spike trimer is accessible only when the RBD is in the standing-up position (18, 20, 21). Hence,

the fact that the MAb binds to the receptor-binding region of the RBD in the spike trimer suggests that the RBD is in the standing-up state. Thus, the results from negative-stain EM and the proteolysis study are consistent with each other, supporting the idea that like DPP4, Mersmab1 stabilizes the RBD in the standing-up position and triggers a conformational change of the spike. Future study on the high-resolution cryo-EM structure of MERS-CoV S-e trimer complexed with Mersmab1 will be needed to provide detailed structural information for the Mersmab1-triggered conformational changes of MERS-CoV S-e.

To understand the pathways of Mersmab1-dependent MERS-CoV entry, we evaluated the potential impact of different proteases on MERS-CoV pseudovirus entry; these proteases are distributed along the viral entry pathway. First, proprotein convertase inhibitor (PPCi) was used for examining the role of proprotein convertases in the maturation of MERS-CoV spike and the impact of proprotein convertases on the ensuing Mersmab1-dependent viral entry (Fig. 5A). The results showed that when MERS-CoV pseudoviruses were produced from HEK293T cells in the presence of PPCi, the cleavage of MERS-CoV spike by proprotein convertases was significantly inhibited (Fig. 5B). In the absence of Mersmab1, MERS-CoV pseudoviruses packaged in the presence of PPCi entered DPP4-expressing human cells more efficiently than those packaged in the absence of PPCi (Fig. 5A). In the presence of Mersmab1, MERS-CoV pseudoviruses packaged in the presence of PPCi entered CD32A-expressing cells more efficiently than those packaged in the absence of PPCi (Fig. 5A). These data suggest that proprotein convertases play a role (albeit not as drastic as some other proteases; see below) in both DPP4-dependent and Mersmab1-dependent MERS-CoV entries. Second, cell surface protease TMPRSS2 (transmembrane serine protease 2) was introduced to human cells for evaluation of its role in Mersmab1-dependent viral entry (Fig. 5C). The results showed that in the absence of Mersmab1, TMPRSS2 enhanced MERS-CoV pseudovirus entry into DPP4-expressing cells, consistent with previous reports (36). In the presence of Mersmab1, TMPRSS2 also enhanced MERS-CoV pseudovirus entry into CD32A-expressing cells, suggesting that TMPRSS2 activates Mersmab1-dependent MERS-CoV entry. Third, lysosomal protease inhibitors were evaluated for the role of lysosomal proteases in Mersmab1-dependent viral entry (Fig. 5D). Two inhibitors were used, lysosomal acidification inhibitor Baf-A1 and cysteine protease inhibitor E64d. The results showed that lysosomal protease inhibitors blocked the DPP4-dependent viral entry pathway, consistent with previous reports (39). Lysosomal protease inhibitors also blocked the Mersmab1-dependent viral entry pathway, suggesting that lysosomal proteases play important roles in Mersmab1-dependent MERS-CoV entry. Taken together, the DPP4-dependent and Mersmab1-dependent MERS-CoV entries can both be activated by proprotein convertases, cell surface proteases, and lysosomal proteases; hence, the same pathways are shared by DPP4-dependent and Mersmab1-dependent MERS-CoV entries.

Antibody dosages for antibody-dependent enhancement of coronavirus entry. To determine the range of Mersmab1 dosages in ADE, MERS-CoV pseudovirus entry was performed in the presence of different concentrations of Mersmab1. Three types of human HEK293T cells were used: HEK293T cells exogenously expressing DPP4 only, CD32A only, or both DPP4 and CD32A. Accordingly, three different sets of results were obtained. First, as the amount of Mersmab1 increased, viral entry into DPP4-expressing HEK293T cells continuously dropped (Fig. 6A). This result reveals that Mersmab1 blocks the DPP4-dependent viral entry pathway by outcompeting DPP4 for binding to MERS-CoV spike. Second, as the amount of Mersmab1 increased, viral entry into CD32A-expressing HEK293T cells first increased and then decreased (Fig. 6A). The turning point was about 100 ng/ml of Mersmab1. A likely explanation for this result is as follows: at low concentrations, more MAb molecules enhance the indirect interactions between MERS-CoV spike and the Fc receptor; at high concentrations, MAb molecules saturate the cell surface Fc receptor molecules and then further bind to MERS-CoV spike and block the indirect interactions between MERS-CoV spike and the Fc receptor. Third, as the amount of Mersmab1 increased, viral entry into cells expressing both DPP4 and

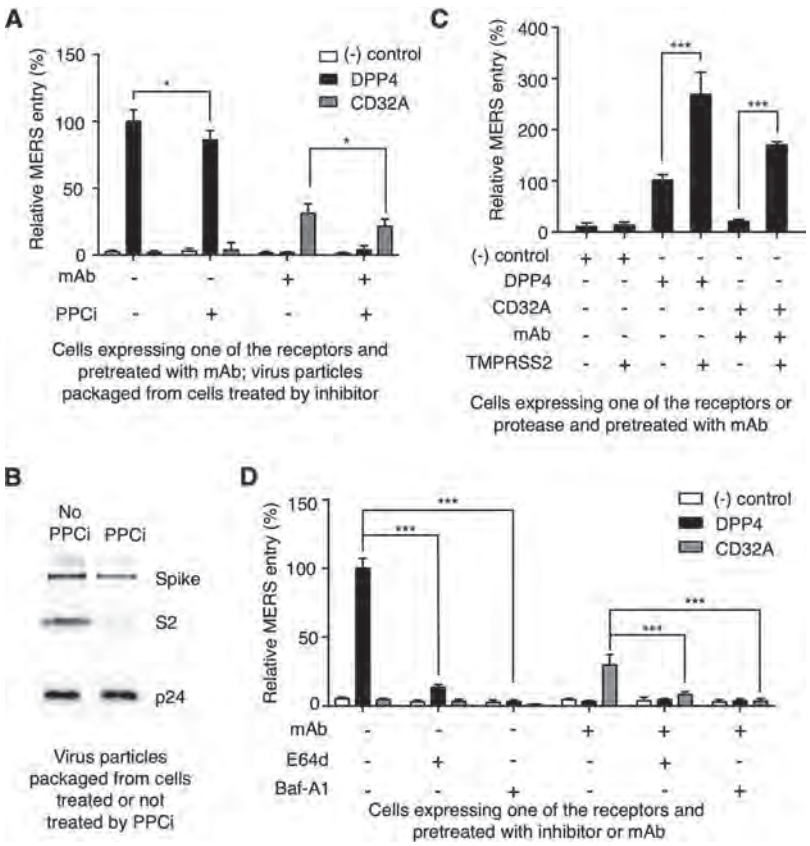


FIG 5 Pathways for antibody-dependent enhancement of coronavirus entry. (A) Impact of proprotein convertases on ADE of MERS-CoV entry. During packaging of MERS-CoV pseudoviruses, HEK293T cells were treated with proprotein convertase inhibitor (PPCi). The MERS-CoV pseudoviruses packaged in the presence of PPCi were then subjected to MERS-CoV pseudovirus entry into HEK293T cells expressing either DPP4 receptor or CD32A receptor. (B) Western blot of MERS-CoV pseudoviruses packaged in the presence or absence of PPCi. MERS-CoV spike protein was detected using anti-C9 antibody targeting its C-terminal C9 tag. As an internal control, another viral protein, p24, was detected using an anti-p24 antibody. (C) Impact of cell surface proteases on ADE of MERS-CoV entry. HEK293T cells exogenously expressing TMPRSS2 (which is a common cell surface protease) were subjected to MERS-CoV pseudovirus entry. TMPRSS2 enhanced both the DPP4-dependent and antibody-dependent entry pathways. (D) Impact of lysosomal proteases on ADE of MERS-CoV entry. HEK293T cells exogenously expressing DPP4 or CD32A were pretreated with one of the lysosomal protease inhibitors E64d and Baf-A1 and then subjected to MERS-CoV pseudovirus entry. Lysosomal protease inhibitors blocked both the DPP4-dependent and antibody-dependent entry pathways. HEK293T cells not expressing DPP4 or CD32A were used as a negative control. All of the experiments were repeated at least three times, with similar results, and representative results are shown. Error bars indicate SD ($n = 4$). Statistical analyses were performed as a one-tailed t test. ***, $P < 0.001$; *, $P < 0.05$. Antibody-dependent and DPP4-dependent viral entries share the same pathways.

CD32A first dropped, then increased, and finally dropped again (Fig. 6B). This result is the cumulative effect of the previous two results. It reveals that when both DPP4 and CD32A are present on host cell surface, Mersmab1 inhibits viral entry (by blocking the DPP4-dependent entry pathway) at low concentrations, promotes viral entry (by enhancing the CD32A-dependent entry pathway) at intermediate concentrations, and inhibits viral entry (by blocking both the DPP4- and CD32A-dependent entry pathways) at high concentrations. We further confirmed the above-described results using MRC5 cells, which are human lung cells endogenously expressing DPP4 (Fig. 6C and D). Therefore, ADE of MERS-CoV entry depends on the range of Mersmab1 dosages as well as expressions of the viral and Fc receptors on cell surfaces.

DISCUSSION

ADE of viral entry has been observed and studied extensively in flaviviruses, particularly dengue virus (3–6). It has also been observed in HIV and Ebola viruses

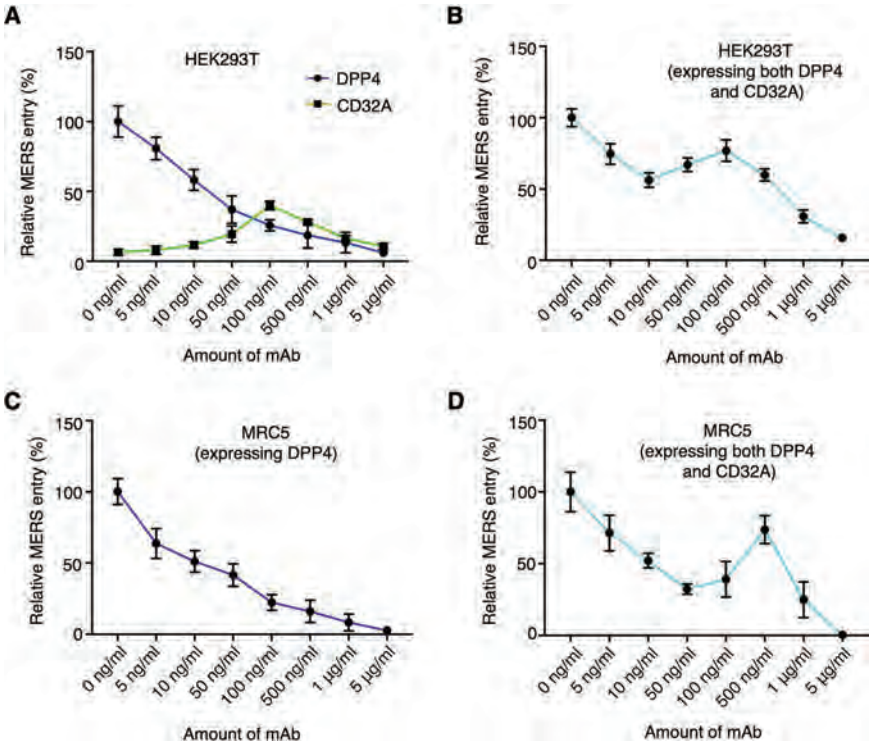


FIG 6 Antibody dosages for antibody-dependent enhancement of coronavirus entry. (A) Impact of antibody dosages on MERS-CoV pseudovirus entry into HEK293T cells exogenously expressing either DPP4 or CD32A. MAb blocks the DPP4-dependent entry pathway; it enhances the antibody-dependent entry pathway at lower concentrations and blocks it at higher concentrations. (B) Impact of antibody dosages on MERS-CoV pseudovirus entry into HEK293T cells exogenously expressing both DPP4 and CD32A. In the presence of both DPP4 and CD32A, MAb blocks viral entry at low concentrations, enhances viral entry at intermediate concentrations, and blocks viral entry at high concentrations. (C) Same experiment as in panel A, except that MRC5 cells replaced HEK293T cells. Here MRC5 cells express DPP4 receptor endogenously. (D) Same experiment as in panel B, except that MRC5 cells replaced HEK293T cells. Here MRC5 cells endogenously express DPP4 and exogenously express CD32A. Please refer to the text for more detailed explanations. All of the experiments were repeated at least three times, with similar results, and representative results are shown. Error bars indicate SD ($n = 4$).

(7–10). For these viruses, it has been proposed that primary viral infections of hosts led to production of antibodies that are subneutralizing or nonneutralizing for secondary viral infections; these antibodies cannot completely neutralize secondary viral infections but instead guide virus particles to enter Fc receptor-expressing cells. ADE can lead to worsened symptoms in secondary viral infections, causing major concerns for epidemiology. ADE is also a major concern for vaccine design and antibody-based drug therapy, since antibodies generated or used in these procedures may lead to ADE. ADE has been observed in coronaviruses for decades, but the molecular mechanisms are unknown. Recent advances in understanding of the receptor recognition and cell entry mechanisms of coronaviruses have allowed us to use coronaviruses as a model system for studying ADE.

In this study, we first demonstrated that a MERS-CoV RBD-specific neutralizing MAb binds to the RBD region of MERS-CoV spike and further showed that the MAb mediates MERS-CoV pseudovirus entry into Fc receptor-expressing human cells. Moreover, a SARS-CoV RBD-specific neutralizing MAb mediates ADE of SARS-CoV pseudovirus entry. These results demonstrated that ADE of coronaviruses is mediated by neutralizing MAbs that target the RBD of coronavirus spikes. In addition, the same coronavirus strains that led to the production of fully neutralizing MAbs can be mediated to go through ADE by these neutralizing MAbs. Our results differ from previously observed ADE of flaviviruses in which primary infections and secondary infections are caused by two different viral strains and in which ADE-mediating MAbs are only subneutralizing

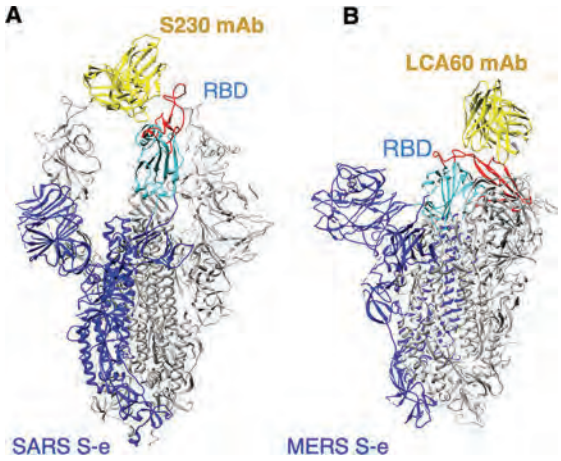


FIG 7 Two previously published structures of coronavirus spike proteins complexed with antibody. (A) SARS-CoV S-e complexed with S230 MAb (PDB code 6NB7). The antibody binds to the side of the RBD, away from the viral-receptor-binding site, stabilizes the RBD in the lying-down state, and hence does not trigger conformational changes of SARS-CoV S-e. (B) MERS-CoV S-e complexed with LCA60 MAb (PDB code 6NB4). The antibody binds to the viral-receptor-binding site in the RBD, stabilizes the RBD in the standing-up state, and hence triggers conformational changes of MERS-CoV S-e.

or nonneutralizing for secondary viral infections (3–6). Therefore, our study expands the concept of ADE of viral entry.

We then examined the molecular mechanism for ADE of coronavirus entry. We showed that the MAb binds to the tip of MERS-CoV spike trimer, where the RBD is located. MAb binding likely stabilizes the RBD in the standing-up position, triggers a conformational change of MERS-CoV spike, and exposes the previously inaccessible S2′ site to proteases. During the preparation of the manuscript, a newly published study demonstrated that a SARS-CoV RBD-specific MAb (named S230) bound to the ACE2-binding region in SARS-CoV RBD, stabilized the RBD in the standing-up position, and triggered conformational changes of SARS-CoV spike (Fig. 7A) (52). In contrast, a MERS-CoV RBD-specific MAb (named LCA60) bound to the side of MERS-CoV RBD, away from the DPP4-binding region, stabilized the RBD in the lying-down position, and did not trigger conformational changes of MERS-CoV spike (Fig. 7B). These published results are consistent with our results regarding Mersmab1-triggered conformational changes of MERS-CoV spike, together suggesting that in order to trigger conformational changes of coronavirus spikes, MAbs need to bind to the receptor-binding region in their RBD and stabilize the RBD in the standing-up position. Moreover, our study revealed that ADE of MERS-CoV entry follows the same entry pathways as DPP4-dependent MERS-CoV entry. Specifically, proprotein convertases partially activate MERS-CoV spike. If cell surface proteases are present, MERS-CoV spike can be further activated and fuse membranes on the cell surface; otherwise, MERS-CoV enters endosomes and lysosomes, where lysosomal proteases activate MERS-CoV spike for membrane fusion. Taken together, our results show that RBD-specific neutralizing MAbs bind to the same region on coronavirus spikes as viral receptors do, trigger conformational changes of the spikes as viral receptors do, and mediate ADE through the same pathways as viral-receptor-dependent viral entry. In other words, RBD-specific neutralizing MAbs mediate ADE of coronavirus entry by functionally mimicking viral receptors.

Finally, we analyzed ADE of coronavirus entry at different antibody dosages. MERS-CoV entry into cells expressing both viral and Fc receptors demonstrates complex MAb-dosage-dependent patterns. As the concentration of MAb increases, (i) viral entry into DPP4-expressing cells is inhibited more efficiently because MAb binds to the spike and blocks the DPP4-dependent entry pathway, (ii) viral entry into Fc receptor-expressing cells is first enhanced and then inhibited because MAb binds to the Fc receptor to enhance the ADE pathway until the Fc receptor molecules are saturated, and (iii) viral entry into cells

expressing both DPP4 and Fc receptor is first inhibited, then enhanced, and finally inhibited again because of the cumulative effects of the previous two patterns. In other words, for viral entry into cells expressing both DPP4 and Fc receptor, there exists a balance between the DPP4-dependent and antibody-dependent entry pathways that can be shifted and determined by MAb dosages. Importantly, ADE occurs only at intermediate MAb dosages. Our study explains an earlier observation that ADE of dengue viruses occurs only at certain concentrations of MAb (5). While many human tissues express either DPP4 or Fc receptor, a few of them, most notably placenta, express both of them (53, 54). For other viruses that use viral receptors different from DPP4, there may also be human tissues where the viral receptor and Fc receptor are both expressed. The expression levels of these two receptors in specific tissue cells likely are determinants of MAb dosages at which ADE would occur in these tissues. Other determinants of ADE-enabling MAb dosages may include the binding affinities of the MAb for the viral and Fc receptors. Overall, our study suggests that ADE of viruses depends on antibody dosages, tissue-specific expressions of viral and Fc receptors, and some intrinsic features of the antibody.

Our findings not only reveal a novel molecular mechanism for ADE of coronaviruses but also provide general guidelines on viral vaccine design and antibody-based antiviral drug therapy. As we have shown here, RBD-specific neutralizing MAbs may mediate ADE of viruses by mimicking the functions of viral receptors. Neutralizing MAbs targeting other parts of viral spikes would be less likely to mediate ADE if they do not trigger the conformational changes of the spikes. Hence, to reduce the likelihood of ADE, spike-based subunit vaccines lacking the RBD can be designed to prevent viral infections. Based on the same principle, neutralizing MAbs targeting other parts of the spike can be selected to treat viral infections. Moreover, as already discussed, our study stresses on the importance of choosing antibody dosages that do not cause ADE and points out that different tissue cells should be closely monitored for potential ADE at certain antibody dosages.

The *in vitro* systems used in this study provide a model framework for ADE. Future research using *in vivo* systems is needed to further confirm these results. Our previous study showed that a humanized version of Mersmab1 efficiently protected human DPP4-transgenic mice from live MERS-CoV challenges (48, 55), suggesting that given the antibody dosages used in this previous study as well as the binding affinity of the MAb for human DPP4, the receptor-dependent pathway of MERS-CoV entry dominated over ADE *in vivo*. Thus, future *in vivo* studies may need to screen for a wide range of antibody dosages and also for a variety of tissues with different ratios of DPP4 and Fc receptor expressions. Although ADE has not been observed for MERS-CoV *in vivo*, our study suggests that ADE occurs under some specific conditions *in vivo*, depending on the antibody dosages, binding affinity of the MAb for DPP4, and tissue expressions of DPP4 and Fc receptor. Moreover, the mechanism that we have identified for ADE of MERS-CoV *in vitro* may account for the ADE observed *in vivo* for other coronaviruses, such as SARS-CoV and feline coronavirus (42–47). Overall, our study reveals complex roles of antibodies in viral entry and can guide future vaccine design and antibody-based drug therapy.

MATERIALS AND METHODS

Cell lines and plasmids. HEK293T cells and HEK293F cells (human embryonic kidney cells), HeLa cells (human cervical cells), and MRC5 cells (human lung cells) were obtained from the American Type Culture Collection (ATCC). HEK293-gamma chain cells (human embryonic kidney cells) were constructed previously (56). These cells were cultured in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine, 100 U/ml of penicillin, and 100 μg/ml of streptomycin. THP-1 cells (human macrophages) were obtained from the ATCC and were cultured in RPMI culture medium (Invitrogen) containing 10% heat-inactivated FBS and supplemented with 10 mM HEPES, 1 mM pyruvate, 2.5 g/liter of D-glucose, 50 pM β-mercaptoethanol, and 100 μg/ml of streptomycin.

For induction of macrophages, human monocytic THP-1 cells were treated with 150 nM phorbol 12-myristate 13-acetate for 24 h, followed by 24 h of incubation in RPMI medium (57) before experiments.

The full-length genes of MERS-CoV spike (GenBank accession number AFS88936.1), SARS-CoV spike (GenBank accession number AFR58742), human DPP4 (GenBank accession number NM_001935.4), and human ACE2 (GenBank accession number NM_001371415.1) were synthesized (GenScript Biotech). Three

Fc receptor genes, human CD16A (GenBank accession number NM_000569.7), human CD32A (GenBank accession number NM_001136219.3), and human CD64A (GenBank accession number NM_000566.3), were cloned previously (58, 59). For protein expressions on cell surfaces or pseudovirus surfaces, the above-named genes were subcloned into the pcDNA3.1(+) vector (Life Technologies) with a C-terminal C9 tag.

Protein purification and antibody preparation. For ELISA and negative-stain electron microscopic study, recombinant MERS-CoV spike ectodomain (S-e) was prepared. The MERS-CoV S-e (residues 1 to 1294) was subcloned into pCMV vector; it contained a C-terminal GCN4 trimerization tag and a His_6 tag. To stabilize S-e in the prefusion conformation, we followed the procedure from a previous study by introducing mutations to the S1/S2 protease cleavage site (RSVR748-751ASVA) and the S2 region (V1060P and L1061P) (21). MERS-CoV S-e was expressed in HEK293F cells using a FreeStyle 293 mammalian cell expression system (Life Technologies). Briefly, HEK293F cells were transfected with the plasmid encoding MERS-CoV S-e and cultured for 3 days. The protein was harvested from the cell culture medium, purified sequentially on a nickel-nitrilotriacetic aid (Ni-NTA) column and Superdex200 gel filtration column (GE Healthcare), and stored in a buffer containing 20 mM Tris (pH 7.2) and 200 mM NaCl. The ectodomain of human DPP4 was expressed and purified as previously described (39). Briefly, DPP4 ectodomain (residues 39 to 766) containing an N-terminal human CD5 signal peptide and a C-terminal His_6 tag was expressed in insect cells using the Bac-to-Bac expression system (Life Technologies), secreted to cell culture medium, and purified in the same way as MERS-CoV S-e.

Both the MERS-CoV RBD-specific MAb (i.e., Mersmab1) and SARS-CoV RBD-specific MAb (i.e., 33G4) were purified as previously described (48, 49). Briefly, hybridoma cells expressing the MAb were injected into the abdomens of mice. After 7 to 10 days, the mouse ascites containing the MAb were collected. The MAb was then purified using a protein A column (GE Healthcare). Fab of Mersmab1 antibody was prepared using immobilized papain beads (Thermo Fisher Scientific) according to the manufacturer's manual. Briefly, Mersmab1 antibody was incubated with immobilized papain beads in digestion buffer (20 mM sodium phosphate, 10 mM EDTA, 20 mM L-cysteine HCl [pH 7.0]) in a shaking water bath at 37°C overnight. After digestion, the reaction was stopped with 10 mM Tris HCl (pH 7.5), and the supernatant was collected through centrifugation at 12,000 × *g* for 15 min. Fab was then separated from undigested IgG and Fc using a protein A column (GE HealthCare).

ELISA. The binding affinity between MAb and MERS-CoV S-e or RBD was measured using ELISA as previously described (60). Briefly, ELISA plates were precoated with MAb (350 nM) at 37°C for 1 h. After blocking with 1% bovine serum albumin (BSA) at 37°C for 1 h, MERS-CoV S-e or RBD (300 nM or gradient concentrations as specified in Fig. 2) was added to the plates and incubated with MAb at 37°C for 1 h. After washes with phosphate-buffered saline (PBS), the plates were incubated with anti-His_6 antibody (Santa Cruz) at 37°C for 1 h. Then the plates were washed with PBS and incubated with horseradish peroxidase (HRP)-conjugated goat anti-mouse IgG antibody (1:5,000) at 37°C for 1 h. After more washes with PBS, the enzymatic reaction was carried out using ELISA substrate (Life Technologies) and stopped with 1 M H_2SO_4. Absorbance at 450 nm (A_{450}) was measured using a Tecan Infinite M1000 PRO microplate reader (Tecan Group Ltd.). Five replicates were done for each sample. PBS was used as a negative control.

Flow cytometry cell-binding assay. Flow cytometry was performed as previously described (22). Briefly, HEK293T cells exogenously expressing DPP4 or one of the Fc receptors were incubated with MERS-CoV S-e (40 μg/ml) and MAb (50 μg/ml) (both of which contained a C-terminal His_6 tag) at room temperature for 30 min, followed by incubation with fluorescein phycoerythrin (PE)-labeled anti-His_6 probe antibody for another 30 min. The cells then were analyzed using fluorescence-activated cell sorting (FACS).

Pseudovirus entry assay. The coronavirus spike-mediated pseudovirus entry assay was carried out as previously described (61, 62). Briefly, for pseudovirus packaging, HEK293T cells were cotransfected with a plasmid carrying an Env-defective, luciferase-expressing HIV type 1 genome (pNL4–3.luc.R-E-) and a plasmid encoding MERS-CoV or SARS-CoV spike. Pseudoviruses were harvested and purified using a sucrose gradient ultracentrifugation at 40,000 × *g* 72 h after transfection and then used to enter the target cells. To detect pseudovirus entry, pseudoviruses and cells were incubated for 5 h at 37°C, and then medium was changed and cells were incubated for an additional 60 h. Cells were then washed with PBS and lysed. Aliquots of cell lysates were transferred to Optiplate-96 (PerkinElmer), followed by addition of luciferase substrate. Relative light unites (RLUs) were measured using an EnSpire plate reader (PerkinElmer). All the measurements were carried out in four replicates. To inhibit proprotein convertases during packaging of MERS-CoV pseudoviruses, 50 nM proprotein convertase inhibitor (PPCi) Dec-RVKR-CMK (Enzo Life Sciences) was added to the cell culture medium 5 h posttransfection, before the packaged pseudoviruses were purified as described above. Inhibition of pseudovirus entry using various protease inhibitors was carried out as described previously (63). Briefly, cells were pretreated with 50 nM proprotein convertase inhibitor Dec-RVKR-CMK (Enzo Life Sciences), 100 nM camostat mesylate (Sigma-Aldrich), 100 nM bafilomycin A1 (Baf-A1) (Sigma-Aldrich), and 50 nM E64d (Sigma-Aldrich) at 37°C for 1 h or 500 ng/ml antibody for 5 min. The above-described cells were then used for pseudovirus entry assay.

Isolation and quantification of cell surface receptor proteins. To examine the expression levels of receptor proteins in cell membranes, the cells expressing the receptor were harvested and all membrane-associated proteins were extracted using a membrane protein extraction kit (Thermo Fisher Scientific). Briefly, cells were centrifuged at 300 × *g* for 5 min and washed with cell wash solution twice. The cell pellets were resuspended in 0.75 ml of permeabilization buffer and incubated at 4°C for 10 min. The supernatant containing cytosolic proteins was removed after centrifugation at 16,000 × *g* for 15 min. The pellets containing membrane-associated proteins were resuspended in 0.5 ml of solubilization buffer and incubated at 4°C for 30 min. After centrifugation at 16,000 × *g* for 15 min, the membrane-associated

proteins from the supernatant were transferred to a new tube. The expression level of membrane-associated C9-tagged receptor proteins among all membrane-associated proteins was then measured using Western blot analysis and further used for normalizing the results from flow cytometry cell-binding assays and pseudovirus entry assays.

Extraction of total RNA and qRT-PCR. Total RNAs of cells were extracted using TRIzol reagent according to the manufacturer's manual. Briefly, TRIzol was added to the cell lysate, and then chloroform and phenol-chloroform were added to precipitate RNA. The RNA pellets were washed using ethanol, solubilized in diethyl pyrocarbonate (DEPC)-treated water, and then reverse transcribed using murine leukemia virus (MLV) reverse transcriptase (Promega) and oligo(dT) primers (Promega). Quantitative PCR on DPP4 RNA was performed using DPP4-specific primers and a SYBR qPCR kit (Bio-Rad) in a CFX qPCR instrument (Bio-Rad). Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) RNA was used as a control. The primers are as follows: for DPP4, forward, 5′-AGTGGCGTGTTCAAGTGTGG-3′, and reverse, 5′-CAAG GTTGTCTTCTGGAGTTGG-3′; for GAPDH, forward, 5′-GGAAGGTGAAGGTCGGAGTCAACGG-3′, and reverse, 5′-CTCGCTCCTGGAAGATGGTGATGGG-3′.

Proteolysis assay. Purified MERS-CoV pseudoviruses were incubated with 67 μg/ml of recombinant DPP4, 67 μg/ml of MAb, or PBS at 37°C for 30 min and then treated with 10^{-3} mg/ml of tosylsulfonyl phenylalanyl chloromethyl ketone (TPCK)-treated trypsin on ice for 20 min. Samples were subjected to Western blotting. MERS-CoV spike and its cleaved fragments (which contained a C-terminal C9 tag) were detected using an anti-C9 tag monoclonal antibody (Santa Cruz Biotechnology).

Negative-stain electron microscopy. Samples were diluted to a final concentration of 0.02 mg/ml in PBS and loaded onto glow-discharged 400-mesh carbon grids (Electron Microscopy Sciences). The grids were stained with 0.75% uranyl formate. All micrographs were acquired using a Tecnai G2 Spirit BioTWIN at 120 keV (FEI Company) and an Eagle 4-megapixel charge-coupled-device (CCD) camera at 6,000 × nominal magnification at the University of Minnesota.

ACKNOWLEDGMENTS

We thank Matthew Aliota for comments.

This work was supported by NIH grants R01AI089728 (to F.L), R01AI110700 (to F.L.), and R01AI139092 (to L.D. and F.L.).

REFERENCES

1. Tirado SM, Yoon KJ. 2003. Antibody-dependent enhancement of virus infection and disease. Viral Immunol 16:69–86. https://doi.org/10.1089/088282403763635465.
2. Takada A, Kawaoka Y. 2003. Antibody-dependent enhancement of viral infection: molecular mechanisms and in vivo implications. Rev Med Virol 13:387–398. https://doi.org/10.1002/rmv.405.
3. Guzman MG, Alvarez M, Rodriguez-Roche R, Bernardo L, Montes T, Vazquez S, Morier L, Alvarez A, Gould EA, Kouri G, Halstead SB. 2007. Neutralizing antibodies after infection with dengue 1 virus. Emerg Infect Dis 13:282–286. https://doi.org/10.3201/eid1302.060539.
4. Dejnirattisai W, Jumnainsong A, Onsirisakul N, Fitton P, Vasanawathana S, Limpitikul W, Puttikhunt C, Edwards C, Duangchinda T, Supasa S, Chawansuntati K, Malasit P, Mongkolsapaya J, Screaton G. 2010. Cross-reacting antibodies enhance dengue virus infection in humans. Science 328:745–748. https://doi.org/10.1126/science.1185181.
5. Katzelnick LC, Gresh L, Halloran ME, Mercado JC, Kuan G, Gordon A, Balmaseda A, Harris E. 2017. Antibody-dependent enhancement of severe dengue disease in humans. Science 358:929–932. https://doi.org/10.1126/science.aan6836.
6. Whitehead SS, Blaney JE, Durbin AP, Murphy BR. 2007. Prospects for a dengue virus vaccine. Nat Rev Microbiol 5:518–528. https://doi.org/10.1038/nrmicro1690.
7. Willey S, Aasa-Chapman MMI, O'Farrell S, Pellegrino P, Williams I, Weiss RA, Neil SJD. 2011. Extensive complement-dependent enhancement of HIV-1 by autologous non-neutralising antibodies at early stages of infection. Retrovirology 8:16. https://doi.org/10.1186/1742-4690-8-16.
8. Beck Z, Prohaszka Z, Fust G. 2008. Traitors of the immune system-enhancing antibodies in HIV infection: their possible implication in HIV vaccine development. Vaccine 26:3078–3085. https://doi.org/10.1016/j.vaccine.2007.12.028.
9. Takada A, Watanabe S, Okazaki K, Kida H, Kawaoka Y. 2001. Infectivity-enhancing antibodies to Ebola virus glycoprotein. J Virol 75:2324–2330. https://doi.org/10.1128/JVI.75.5.2324-2330.2001.
10. Takada A, Feldmann H, Ksiazek TG, Kawaoka Y. 2003. Antibody-dependent enhancement of Ebola virus infection. J Virol 77:7539–7544. https://doi.org/10.1128/jvi.77.13.7539-7544.2003.
11. Perlman S, Netland J. 2009. Coronaviruses post-SARS: update on replication and pathogenesis. Nat Rev Microbiol 7:439–450. https://doi.org/10.1038/nrmicro2147.
12. Enjuanes L, Almazan F, Sola I, Zuniga S. 2006. Biochemical aspects of coronavirus replication and virus-host interaction. Annu Rev Microbiol 60:211–230. https://doi.org/10.1146/annurev.micro.60.080805.142157.
13. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus A, Fouchier R. 2012. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med 367:1814–1820. https://doi.org/10.1056/NEJMoa1211721.
14. Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, Tong SX, Urbani C, Comer JA, Lim W, Rollin PE, Dowell SF, Ling AE, Humphrey CD, Shieh WJ, Guarner J, Paddock CD, Rota P, Fields B, DeRisi J, Yang JY, Cox N, Hughes JM, LeDuc JW, Bellini WJ, Anderson LJ. 2003. A novel coronavirus associated with severe acute respiratory syndrome. N Engl J Med 348:1953–1966. https://doi.org/10.1056/NEJMoa030781.
15. Peiris JSM, Lai ST, Poon LLM, Guan Y, Yam LYC, Lim W, Nicholls J, Yee WKS, Yan WW, Cheung MT, Cheng VCC, Chan KH, Tsang DNC, Yung RWH, Ng TK, Yuen KY. 2003. Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet 361:1319–1325. https://doi.org/10.1016/S0140-6736(03)13077-2.
16. de Groot RJ, Baker SC, Baric RS, Brown CS, Drosten C, Enjuanes L, Fouchier RA, Galiano M, Gorbalenya AE, Memish ZA, Perlman S, Poon LL, Snijder EJ, Stephens GM, Woo PC, Zaki AM, Zambon M, Ziebuhr J. 2013. Middle East respiratory syndrome coronavirus (MERS-CoV): announcement of the Coronavirus Study Group. J Virol 87:7790–7792. https://doi.org/10.1128/JVI.01244-13.
17. Li F. 2016. Structure, function, and evolution of coronavirus spike proteins. Annu Rev Virol 3:237–261. https://doi.org/10.1146/annurev-virology-110615-042301.
18. Yuan Y, Cao D, Zhang Y, Ma J, Qi J, Wang Q, Lu G, Wu Y, Yan J, Shi Y, Zhang X, Gao GF. 2017. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. Nat Commun 8:15092. https://doi.org/10.1038/ncomms15092.
19. Shang J, Zheng Y, Yang Y, Liu C, Geng Q, Luo C, Zhang W, Li F. 2018. Cryo-EM structure of infectious bronchitis coronavirus spike protein reveals structural and functional evolution of coronavirus spike proteins. PLoS Pathog 14:e1007009. https://doi.org/10.1371/journal.ppat.1007009.
20. Song W, Gui M, Wang X, Xiang Y. 2018. Cryo-EM structure of the SARS

coronavirus spike glycoprotein in complex with its host cell receptor ACE2. PLoS Pathog 14:e1007236. https://doi.org/10.1371/journal.ppat.1007236.
21. Pallesen J, Wang N, Corbett KS, Wrapp D, Kirchdoerfer RN, Turner HL, Cottrell CA, Becker MM, Wang L, Shi W, Kong WP, Andres EL, Kettenbach AN, Denison MR, Chappell JD, Graham BS, Ward AB, McLellan JS. 2017. Immunogenicity and structures of a rationally designed prefusion MERS-CoV spike antigen. Proc Natl Acad Sci U S A 114:E7348–E7357. https://doi.org/10.1073/pnas.1707304114.
22. Shang J, Zheng Y, Yang Y, Liu C, Geng Q, Tai W, Du L, Zhou Y, Zhang W, Li F. 2018. Cryo-electron microscopy structure of porcine deltacoronavirus spike protein in the prefusion state. J Virol 92:e01556-17. https://doi.org/10.1128/JVI.01556-17.
23. Walls AC, Tortorici MA, Bosch BJ, Frenz B, Rottier PJ, DiMaio F, Rey FA, Veesler D. 2016. Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. Nature 531:114–117. https://doi.org/10.1038/nature16988.
24. Walls AC, Tortorici MA, Frenz B, Snijder J, Li W, Rey FA, DiMaio F, Bosch BJ, Veesler D. 2016. Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. Nat Struct Mol Biol 23:899–905. https://doi.org/10.1038/nsmb.3293.
25. Kirchdoerfer RN, Cottrell CA, Wang N, Pallesen J, Yassine HM, Turner HL, Corbett KS, Graham BS, McLellan JS, Ward AB. 2016. Pre-fusion structure of a human coronavirus spike protein. Nature 531:118–121. https://doi.org/10.1038/nature17200.
26. Li F. 2015. Receptor recognition mechanisms of coronaviruses: a decade of structural studies. J Virol 89:1954–1964. https://doi.org/10.1128/JVI.02615-14.
27. Li WH, Moore MJ, Vasilieva N, Sui JH, Wong SK, Berne MA, Somasundaran M, Sullivan JL, Luzuriaga K, Greenough TC, Choe H, Farzan M. 2003. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 426:450–454. https://doi.org/10.1038/nature02145.
28. Raj VS, Mou HH, Smits SL, Dekkers DHW, Muller MA, Dijkman R, Muth D, Demmers JAA, Zaki A, Fouchier RAM, Thiel V, Drosten C, Rottier PJM, Osterhaus A, Bosch BJ, Haagmans BL. 2013. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature 495:251–254. https://doi.org/10.1038/nature12005.
29. Li F, Li WH, Farzan M, Harrison SC. 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science 309:1864–1868. https://doi.org/10.1126/science.1116480.
30. Lu G, Hu Y, Wang Q, Qi J, Gao F, Li Y, Zhang Y, Zhang W, Yuan Y, Bao J, Zhang B, Shi Y, Yan J, Gao GF. 2013. Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. Nature 500:227–231. https://doi.org/10.1038/nature12328.
31. Belouzard S, Millet JK, Licitra BN, Whittaker GR. 2012. Mechanisms of coronavirus cell entry mediated by the viral spike protein. Viruses 4:1011–1033. https://doi.org/10.3390/v4061011.
32. Heald-Sargent T, Gallagher T. 2012. Ready, set, fuse! The coronavirus spike protein and acquisition of fusion competence. Viruses 4:557–580. https://doi.org/10.3390/v4040557.
33. Bertram S, Glowacka I, Muller MA, Lavender H, Gnirss K, Nehlmeier I, Niemeyer D, He Y, Simmons G, Drosten C, Soilleux EJ, Jahn O, Steffen I, Pohlmann S. 2011. Cleavage and activation of the severe acute respiratory syndrome coronavirus spike protein by human airway trypsin-like protease. J Virol 85:13363–13372. https://doi.org/10.1128/JVI.05300-11.
34. Matsuyama S, Ujike M, Morikawa S, Tashiro M, Taguchi F. 2005. Protease-mediated enhancement of severe acute respiratory syndrome coronavirus infection. Proc Natl Acad Sci U S A 102:12543–12547. https://doi.org/10.1073/pnas.0503203102.
35. Kam YW, Okumura Y, Kido H, Ng LF, Bruzzone R, Altmeyer R. 2009. Cleavage of the SARS coronavirus spike glycoprotein by airway proteases enhances virus entry into human bronchial epithelial cells in vitro. PLoS One 4:e7870. https://doi.org/10.1371/journal.pone.0007870.
36. Shirato K, Kawase M, Matsuyama S. 2013. Middle East respiratory syndrome coronavirus infection mediated by the transmembrane serine protease TMPRSS2. J Virol 87:12552–12561. https://doi.org/10.1128/JVI.01890-13.
37. Gierer S, Müller MA, Heurich A, Ritz D, Springstein BL, Karsten CB, Schendzielorz A, Gnirß K, Drosten C, Pöhlmann S. 2015. Inhibition of proprotein convertases abrogates processing of the Middle Eastern respiratory syndrome coronavirus spike protein in infected cells but does not reduce viral infectivity. J Infect Dis 211:889–897. https://doi.org/10.1093/infdis/jiu407.
38. Gierer S, Bertram S, Kaup F, Wrensch F, Heurich A, Krämer-Kühl A, Welsch K, Winkler M, Meyer B, Drosten C, Dittmer U, von Hahn T, Simmons G, Hofmann H, Pöhlmann S. 2013. The spike protein of the emerging betacoronavirus EMC uses a novel coronavirus receptor for entry, can be activated by TMPRSS2, and is targeted by neutralizing antibodies. J Virol 87:5502–5511. https://doi.org/10.1128/JVI.00128-13.
39. Zheng Y, Shang J, Yang Y, Liu C, Wan Y, Geng Q, Wang M, Baric R, Li F. 2018. Lysosomal proteases are a determinant of coronavirus tropism. J Virol 92:e01504-18. https://doi.org/10.1128/JVI.01504-18.
40. Walls AC, Tortorici MA, Snijder J, Xiong X, Bosch BJ, Rey FA, Veesler D. 2017. Tectonic conformational changes of a coronavirus spike glycoprotein promote membrane fusion. Proc Natl Acad Sci U S A 114:11157–11162. https://doi.org/10.1073/pnas.1708727114.
41. Li F, Berardi M, Li WH, Farzan M, Dormitzer PR, Harrison SC. 2006. Conformational states of the severe acute respiratory syndrome coronavirus spike protein ectodomain. J Virol 80:6794–6800. https://doi.org/10.1128/JVI.02744-05.
42. Wang SF, Tseng SP, Yen CH, Yang JY, Tsao CH, Shen CW, Chen KH, Liu FT, Liu WT, Chen YM, Huang JC. 2014. Antibody-dependent SARS coronavirus infection is mediated by antibodies against spike proteins. Biochem Biophys Res Commun 451:208–214. https://doi.org/10.1016/j.bbrc.2014.07.090.
43. Kam YW, Kien F, Roberts A, Cheung YC, Lamirande EW, Vogel L, Chu SL, Tse J, Guarner J, Zaki SR, Subbarao K, Peiris M, Nal B, Altmeyer R. 2007. Antibodies against trimeric S glycoprotein protect hamsters against SARS-CoV challenge despite their capacity to mediate FcγRII-dependent entry into B cells in vitro. Vaccine 25:729–740. https://doi.org/10.1016/j.vaccine.2006.08.011.
44. Jaume M, Yip MS, Cheung CY, Leung HL, Li PH, Kien F, Dutry I, Callendret B, Escriou N, Altmeyer R, Nal B, Daeron M, Bruzzone R, Peiris JS. 2011. Anti-severe acute respiratory syndrome coronavirus spike antibodies trigger infection of human immune cells via a pH- and cysteine protease-independent FcγR pathway. J Virol 85:10582–10597. https://doi.org/10.1128/JVI.00671-11.
45. Corapi WV, Olsen CW, Scott FW. 1992. Monoclonal antibody analysis of neutralization and antibody-dependent enhancement of feline infectious peritonitis virus. J Virol 66:6695–6705.
46. Hohdatsu T, Yamada M, Tominaga R, Makino K, Kida K, Koyama H. 1998. Antibody-dependent enhancement of feline infectious peritonitis virus infection in feline alveolar macrophages and human monocyte cell line U937 by serum of cats experimentally or naturally infected with feline coronavirus. J Vet Med Sci 60:49–55. https://doi.org/10.1292/jvms.60.49.
47. Vennema H, de Groot RJ, Harbour DA, Dalderup M, Gruffydd-Jones T, Horzinek MC, Spaan WJ. 1990. Early death after feline infectious peritonitis virus challenge due to recombinant vaccinia virus immunization. J Virol 64:1407–1409.
48. Du L, Zhao G, Yang Y, Qiu H, Wang L, Kou Z, Tao X, Yu H, Sun S, Tseng CT, Jiang S, Li F, Zhou Y. 2014. A conformation-dependent neutralizing monoclonal antibody specifically targeting receptor-binding domain in Middle East respiratory syndrome coronavirus spike protein. J Virol 88:7045–7053. https://doi.org/10.1128/JVI.00433-14.
49. Du L, Zhao G, Li L, He Y, Zhou Y, Zheng BJ, Jiang S. 2009. Antigenicity and immunogenicity of SARS-CoV S protein receptor-binding domain stably expressed in CHO cells. Biochem Biophys Res Commun 384:486–490. https://doi.org/10.1016/j.bbrc.2009.05.003.
50. He YX, Lu H, Siddiqui P, Zhou YS, Jiang SB. 2005. Receptor-binding domain of severe acute respiratory syndrome coronavirus spike protein contains multiple conformation-dependent epitopes that induce highly potent neutralizing antibodies. J Immunol 174:4908–4915. https://doi.org/10.4049/jimmunol.174.8.4908.
51. Millet JK, Whittaker GR. 2014. Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. Proc Natl Acad Sci U S A 111:15214–15219. https://doi.org/10.1073/pnas.1407087111.
52. Walls AC, Xiong X, Park YJ, Tortorici MA, Snijder J, Quispe J, Cameroni E, Gopal R, Dai M, Lanzavecchia A, Zambon M, Rey FA, Corti D, Veesler D. 2019. Unexpected receptor functional mimicry elucidates activation of coronavirus fusion. Cell 176:1026–1039.e1015. https://doi.org/10.1016/j.cell.2018.12.028.
53. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA-K, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist P-H, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J, Pontén F.

2015. Proteomics. Tissue-based map of the human proteome. Science 347:1260419. https://doi.org/10.1126/science.1260419.

54. Uhlén M, Björling E, Agaton C, Szigyarto CA-K, Amini B, Andersen E, Andersson A-C, Angelidou P, Asplund A, Asplund C, Berglund L, Bergström K, Brumer H, Cerjan D, Ekström M, Elobeid A, Eriksson C, Fagerberg L, Falk R, Fall J, Forsberg M, Björklund MG, Gumbel K, Halimi A, Hallin I, Hamsten C, Hansson M, Hedhammar M, Hercules G, Kampf C, Larsson K, Lindskog M, Lodewyckx W, Lund J, Lundeberg J, Magnusson K, Malm E, Nilsson P, Odling J, Oksvold P, Olsson I, Oster E, Ottosson J, Paavilainen L, Persson A, Rimini R, Rockberg J, Runeson M, Sivertsson A, Sköllermo A, Steen J, Stenvall M, Sterky F, Strömberg S, Sundberg M, Tegel H, Tourle S, Wahlund E, Waldén A, Wan J, Wernérus H, Westberg J, Wester K, Wrethagen U, Xu LL, Hober S, Pontén F. 2005. A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics 4:1920–1932. https://doi.org/10.1074/mcp.M500279-MCP200.
55. Qiu H, Sun S, Xiao H, Feng J, Guo Y, Tai W, Wang Y, Du L, Zhao G, Zhou Y. 2016. Single-dose treatment with a humanized neutralizing antibody affords full protection of a human transgenic mouse model from lethal Middle East respiratory syndrome (MERS)-coronavirus infection. Antiviral Res 132:141–148. https://doi.org/10.1016/j.antiviral.2016.06.003.
56. Jing Y, Ni Z, Wu J, Higgins L, Markowski TW, Kaufman DS, Walcheck B. 2015. Identification of an ADAM17 cleavage region in human CD16 (FcgammaRIII) and the engineering of a non-cleavable version of the receptor in NK cells. PLoS One 10:e0121788. https://doi.org/10.1371/journal.pone.0121788.
57. Genin M, Clement F, Fattaccioli A, Raes M, Michiels C. 2015. M1 and M2 macrophages derived from THP-1 cells differentially modulate the response of cancer cells to etoposide. BMC Cancer 15:577. https://doi.org/10.1186/s12885-015-1546-9.
58. Snyder KM, Hullsiek R, Mishra HK, Mendez DC, Li Y, Rogich A, Kaufman DS, Wu J, Walcheck B. 2018. Expression of a recombinant high affinity IgG Fc receptor by engineered NK cells as a docking platform for therapeutic mAbs to target cancer cells. Front Immunol 9:2873. https://doi.org/10.3389/fimmu.2018.02873.
59. Su K, Li X, Edberg JC, Wu J, Ferguson P, Kimberly RP. 2004. A promoter haplotype of the immunoreceptor tyrosine-based inhibitory motif-bearing FcgammaRIIb alters receptor expression and associates with autoimmunity. II. Differential binding of GATA4 and Yin-Yang1 transcription factors and correlated receptor expression and function. J Immunol 172:7192–7199. https://doi.org/10.4049/jimmunol.172.11.7192.
60. Du L, Tai W, Yang Y, Zhao G, Zhu Q, Sun S, Liu C, Tao X, Tseng CK, Perlman S, Jiang S, Zhou Y, Li F. 2016. Introduction of neutralizing immunogenicity index to the rational design of MERS coronavirus subunit vaccines. Nat Commun 7:13473. https://doi.org/10.1038/ncomms13473.
61. Liu C, Tang J, Ma Y, Liang X, Yang Y, Peng G, Qi Q, Jiang S, Li J, Du L, Li F. 2015. Receptor usage and cell entry of porcine epidemic diarrhea coronavirus. J Virol 89:6121–6125. https://doi.org/10.1128/JVI.00430-15.
62. Yang Y, Du L, Liu C, Wang L, Ma C, Tang J, Baric RS, Jiang S, Li F. 2014. Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human transmission of MERS coronavirus. Proc Natl Acad Sci U S A 111:12516–12521. https://doi.org/10.1073/pnas.1405889111.
63. Liu C, Ma Y, Yang Y, Zheng Y, Shang J, Zhou Y, Jiang S, Du L, Li J, Li F. 2016. Cell entry of porcine epidemic diarrhea coronavirus is activated by lysosomal proteases. J Biol Chem 291:24779–24786. https://doi.org/10.1074/jbc.M116.740746.
64. Chen Y, Rajashankar KR, Yang Y, Agnihothram SS, Liu C, Lin YL, Baric RS, Li F. 2013. Crystal structure of the receptor-binding domain from newly emerged Middle East respiratory syndrome coronavirus. J Virol 87:10777–10783. https://doi.org/10.1128/JVI.01756-13.

Exhibit

38

Analysis of hospital traffic and search engine data in Wuhan China indicates early disease activity in the Fall of 2019

Analysis of hospital traffic and search engine data in Wuhan China indicates early disease activity in the Fall of 2019

Elaine Okanyene Nsoesie[1]*, Benjamin Rader[2,3]*, Yiyao L. Barnoon[2], Lauren Goodwin[2], John S. Brownstein[2,4]

1. Department of Global Health, Boston University School of Public Health, Boston, USA
2. Computational Epidemiology Lab, Boston Children's Hospital, Boston, USA
3. Department of Epidemiology, Boston University School of Public Health, Boston, USA
4. Departments of Pediatrics and Biomedical Informatics, Harvard Medical School, Boston, USA

*authors contributed equally
Correspondence should be addressed to: john.brownstein@childrens.harvard.edu

Abstract:
The global COVID-19 pandemic was originally linked to a zoonotic spillover event in Wuhan's Huanan Seafood Market in November or December of 2019. However, recent evidence suggests that the virus may have already been circulating at the time of the outbreak. Here we use previously validated data streams - satellite imagery of hospital parking lots and Baidu search queries of disease related terms - to investigate this possibility. We observe an upward trend in hospital traffic and search volume beginning in late Summer and early Fall 2019. While queries of the respiratory symptom "cough" show seasonal fluctuations coinciding with yearly influenza seasons, "diarrhea" is a more COVID-19 specific symptom and only shows an association with the current epidemic. The increase of both signals precedes the documented start of the COVID-19 pandemic in December, highlighting the value of novel digital sources for surveillance of emerging pathogens.

Introduction

Early investigations into SARS-CoV-2 emergence and the resulting COVID-19 disease outbreak proposed the proximate cause was a zoonotic spillover event in late November or early December 2019 in Wuhan, China[1–3]. This was supported by preliminary epidemiological studies, including the initial clinical series which linked two-thirds of the identified cases to the Huanan Seafood Market in Wuhan[4,5]. Critically, the study found no direct connection to the market for 14 individuals, including the first known case of COVID-19, leaving open the possibility of alternate points of origin and infection[4]. Additionally, virologic samples of wildlife in the Huanan market could not be linked to SARS-CoV-2, suggesting transmission at the market was downstream from the spillover event[6]. Here we consider that SARS-CoV-2 may have already been circulating in the community prior to the identification of the Huanan Market cluster. This hypothesis is supported by emerging epidemiologic and phylogenetic evidence indicating that the virus emerged in southern China[7] and may have already spread internationally[8], and adapted for efficient human transmission[9] by the time it was detected in late December.

Digital epidemiology and non-traditional data streams, such as satellite imagery and internet search trends, have previously been harnessed for respiratory disease surveillance[10,11]. These sources have been shown to be early indicators of epidemics and sensitive to trends that may otherwise go undetected by traditional public health surveillance mechanisms[12,13]. In this study, we use two of these previously validated data streams to look for indicators of potential COVID-19 disease prior to December 2019. First, using vehicle counts extracted from high-resolution satellite imagery of hospital parking lots in Wuhan, we aim to estimate trends in hospital occupancy and its association with influenza-like illness (ILI) trends. This method has been demonstrated as an effective proxy for detecting hospital traffic related to respiratory illness in Latin America[11]. Second, we use Baidu search trends to try to determine the etiology of potential changes in ILI. We have previously shown that the volume of Baidu search queries can be used to estimate influenza trends in China[14]. Together, we assess whether these digital sources can augment traditional epidemiologic, genetic and virologic tools for understanding the timing of SARS-CoV-2 emergence.

Methods

Hospital traffic data

We obtained archived high-resolution satellite imagery (average resolution of about 70 cm) data for Wuhan, China from Remote Sensing Metrics (RS Metrics). We developed a comprehensive list of hospitals in Wuhan (using Google Maps, Wikipedia and PubMed). After the exclusion of sub-specialty hospitals (e.g. Wuhan Asia Heart) and hospitals with no satellite imagery available (Jinyintan), we identified 6 hospitals for imagery analysis: Hubei Women and Children's, Wuhan Tianyou, Wuhan Central, Wuhan Tongji Medical University, Wuhan Union, and Zhongnan Hospital of Wuhan University. We also identified high traffic areas including the Huanan Seafood Market and two railway stations (Wuchang and Hankou) for validation. For each location, RS Metrics identified cars in parking lots by first delineating hospital (or other site) premises, parking lot borders and street parking by automated feature extraction and then manual counting and quality control (as previously described)[11]. Images with tree cover, building

shadow, construction and other factors that present difficulties in defining the contours were excluded since this could lead to over- or under-counting of the number of vehicles. The process of data analysis was independent of the image selection process. The dataset used in analysis consisted of the date and time of each image, the hospital's name and geographic location (including the address, latitude and longitude), and the numbers of vehicles in the parking lot. A metric of relative daily car volume was computed: $relative\ count_{hj} = \frac{raw\ count_{hj}}{mean(raw\ count_{hs})}$, where for each hospital, *h,* and daily satellite image, *j*, counts were compared against baseline means for that hospital during each segment of the week, *s* (weekdays, Saturday, and Sunday). A loess (locally estimated scatterplot smoothing) regression line with a 40% smoothing span was fit to the data.

Search query data

Baidu's database (http://index.baidu.com/) contains logs of web and mobile search query volume in China. User confidentiality is maintained, since only the relative term frequency data is available. We obtained daily data for symptom-related searches likely associated with COVID-19 illness in Wuhan from April 2017 to May 2020. We extracted the relative search volumes of the terms "cough" and "diarrhea" using WebPlotDigitizer, v4.2[15].

Clinical Data

We obtained data on influenza-like illness from two sentinel hospitals in Wuhan: the Children's Hospital of Wuhan and Wuhan No. 1 Hospital, from Kong et al. (2020)[16]. The authors state that ILI trends noted in these two hospitals represent the overall trend in the local population. The two hospitals are the largest pediatric hospital in Hubei and a major general hospital, respectively. Counts of confirmed COVID-19 cases in Wuhan were aggregated from an open access repository of global line-list disease data[17].

Results

We collected 111 satellite images of Wuhan (multiple sites per image) from January 9, 2018 to April 30, 2020 resulting in 140 successful daily extractions of parking lot volume from hospitals (**Figure 1,** example on top panel) and 117 from the three high-volume control sites (**Figure 1,** example on bottom panel). Between 2018 and 2020, there was a general upward trend of increased hospital occupancy as measured by the parking lot volume proxy (**Figure 2, a**). The loess smoothed line shows a steep increase in volume starting in August 2019 (**Figure 2,** first annotation) and culminating with a peak in December 2019 (**Figure 2,** second annotation). Individual hospitals have days of high relative volume in both Fall and Winter 2019. However, between September and October 2019, 5 of the 6 hospitals show their highest relative daily volume of the analyzed series, coinciding with elevated levels of Baidu search queries for the terms "diarrhea" and "cough" (**Figure 2, b**). While searches for "diarrhea" only show elevated traffic starting in late 2019, "cough" shows yearly peaks that approximately coincide with influenza season (**Figure 2, c, turquoise**). Both search query terms show a large increase approximately 3 weeks preceding the large spike of confirmed COVID-19 cases in early 2020 (**Figure 2, c, purple**). There is a large decrease in hospital volume and search query data following the public health lockdown of Wuhan on January 23, 2020 (**Figure 2,** third annotation).

This decrease was seen in both hospital and control sites (**Figure 1**). In Spring 2020, hospital volume begins to trend upward again. In late May 2020 there is a small uptick in Baidu search volume for both “diarrhea” and “cough” in Wuhan.

Discussion

Here we show increased hospital traffic and symptom search data in Wuhan preceded the documented start of the SARS-CoV-2 pandemic in December 2019. While we cannot confirm if the increased volume was directly related to the new virus, our evidence supports other recent work showing that emergence happened before identification at the Huanan Seafood market. These findings also corroborate the hypothesis that the virus emerged naturally in southern China and was potentially already circulating at the time of the Wuhan cluster[7].

In August, we identify a unique increase in searches for diarrhea which was neither seen in previous flu seasons or mirrored in the cough search data. While surprising, this finding lines up with the recent recognition that gastrointestinal (GI) symptoms are a unique feature of COVID-19 disease and may be the chief complaint of a significant proportion of presenting patients [18]. This symptom search increase is then followed by a rise in hospital parking lot traffic in October and November, as well as a rise in searches for cough. While we cannot conclude the reason for this increase, we hypothesize that broad community transmission may have led to more acute cases requiring medical attention, resulting in higher viral loads and worse symptoms[19]. This temporal progression of clinical presentation from mild illness to more severe outcomes has been shown elsewhere[20]. Interestingly, a retrospective study was conducted in Wuhan, China at a hospital designated for the management of patients with COVID-19, which also happens to be represented in our dataset (Wuhan Union Hospital, Wuhan Tongji Medical University)[21]. While respiratory symptoms are common indicators of SARS-CoV-2 infection, this study revealed that a potentially large segment of patients with mainly digestive symptoms, such as diarrhea, may play an important role in community transmission.

The initial rise in GI symptoms may also hint at the missed early signals of COVID-19 in current surveillance systems for respiratory pathogens. The standard definition for influenza-like illness is a combination of fever along with cough and/or sore throat. This narrow definition, which has focused on detection of influenza transmission, would have missed milder cases with a different symptom mix that also could include loss of taste and smell. This finding also hints at the need to broaden surveillance efforts to consider novel pathogens that might display a range of unexpected symptoms. Furthermore, the recent uptick in hospital traffic and search engine query data in May coincides with recent reports of new case clusters in Wuhan[22,23].

The use of satellite data does come with limitations that are amplified in densely populated urban areas. The presence of tall buildings cast shadows that block the view of parking lots, requiring images to be taken at noon local time and exactly overhead. Wuhan experienced a significant amount of cloudy weather during November to February 2019 which along with the consistent smog, created limitations in high quality images that could be harvested. We also experienced challenges in acquiring data from Chinese satellite companies. Finally, prior to the

SARS-CoV-2 outbreak, there are relatively limited archived images of Wuhan compared to other urban centers because of lack of commercial interest.

There are also several limitations to using search query data. We are unable to know the intention of a search and not all symptom searches are necessarily linked to disease morbidity. Search queries resulting from panic and media attention have been shown previously[24] and may have driven the symptom spike we see in January. These data are also vulnerable to fluctuations related to events we might not be aware of and individual search behavior changes over time, which may result in spurious signals[25]. Surveillance using web-query data depends on adequate Internet access and Internet penetration in China can be highly variable. However, by the end of 2017, the internet penetration rate was 70.7% in Wuhan which was 14.9% higher than the national average[26]. Additionally, the use of an automated tool to digitize images does mean the resulting times and values extracted are approximate, although this method has been shown to be effective at replicating time-series[27]. While Baidu query data has been validated for influenza epidemic surveillance[14], investigation of the viability of these data to monitor COVID-19 is still in progress[28]. Nonetheless, when looking at Google search query for the United States, we see a similar pattern of rising GI and cough symptoms alongside confirmed cases.

Our retrospective analysis cannot verify if increased hospital and search engine volume is related to the SARS-CoV-2 virus. While alternative explanations such as the 2019 Military World Games in Wuhan may explain some increases in parking lot traffic, this event opened on October 18, 2019, weeks after the initial rise in Baidu search engine traffic. Still, further research is needed to validate the emergence of SARS-CoV-2. This study adds to a growing body of work on the value of digital sources as an early indicator of a disease outbreak in the context of limited integrated electronic surveillance data. While not a replacement for more traditional methods, these data can help supplement other sources to provide a richer situational awareness picture of social disruptions, including those caused by a novel respiratory virus.

Author contributions: EON and JSB contributed to conceptualization. YB, LG and JSB contributed to data acquisition. EON, BR and JSB contributed to data analysis. BR and JSB wrote the first draft of the manuscript. All authors contributed to interpretation of results and final manuscript writing. All authors have seen and approved the manuscript.

Competing interests: The authors did not receive funding for this work and have no conflicts of interest to declare.

References

1. Lu R, Zhao X, Li J, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*. 2020;395(10224):565-574. doi:10.1016/S0140-6736(20)30251-8
2. Benvenuto D, Giovanetti M, Ciccozzi A, Spoto S, Angeletti S, Ciccozzi M. The 2019-new coronavirus epidemic: Evidence for virus evolution. *J Med Virol*. 2020;92(4):455-459. doi:10.1002/jmv.25688
3. Duchene S, Featherstone L, Haritopoulou-Sinanidou M, Rambaut A, Lemey P, Baele G. Temporal signal and the phylodynamic threshold of SARS-CoV-2. *bioRxiv*. May 2020:2020.05.04.077735. doi:10.1101/2020.05.04.077735
4. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*. 2020;395(10223):497-506. doi:10.1016/S0140-6736(20)30183-5
5. Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med*. 2020;382(8):727-733. doi:10.1056/NEJMoa2001017
6. Wuhan's Huanan seafood market a victim of COVID-19: CDC director - Global Times. https://www.globaltimes.cn/content/1189506.shtml. Accessed June 1, 2020.
7. Latinne A, Hu B, Olival KJ, et al. Origin and cross-species transmission of bat coronaviruses in China. *bioRxiv*. May 2020:2020.05.31.116061. doi:10.1101/2020.05.31.116061
8. Deslandes A, Berti V, Tandjaoui-Lambotte Y, et al. SARS-COV-2 was already spreading in France in late December 2019. *Int J Antimicrob Agents*. May 2020:106006. doi:10.1016/j.ijantimicag.2020.106006
9. Zhan SH, Deverman BE, Chan YA. SARS-CoV-2 is well adapted for humans. What does this mean for re-emergence? *bioRxiv*. May 2020:2020.05.01.073262. doi:10.1101/2020.05.01.073262
10. Salathé M, Freifeld CC, Mekaru SR, Tomasulo AF, Brownstein JS. Influenza A (H7N9) and the importance of digital epidemiology. *N Engl J Med*. 2013;369(5):401-404. doi:10.1056/NEJMp1307752
11. Nsoesie EO, Butler P, Ramakrishnan N, Mekaru SR, Brownstein JS. Monitoring Disease Trends using Hospital Traffic Data from High Resolution Satellite Imagery: A Feasibility Study. *Sci Rep*. 2015;5(1):1-8. doi:10.1038/srep09112
12. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection - Harnessing the web for public health surveillance. *N Engl J Med*. 2009;360(21):2153-2157. doi:10.1056/NEJMp0900702
13. Dai Y, Wang J. Identification of COVID-19 Outbreak Signals Prior to the Traditional Disease Surveillance System. *SSRN Electron J*. May 2020. doi:10.2139/ssrn.3578808
14. Yuan Q, Nsoesie EO, Lv B, Peng G, Chunara R, Brownstein JS. Monitoring Influenza Epidemics in China with Search Query from Baidu. Cowling BJ, ed. *PLoS One*. 2013;8(5):e64323. doi:10.1371/journal.pone.0064323
15. Rohatgi A. WebPlotDigitizer. 2019. https://automeris.io/WebPlotDigitizer/blog/index.html. Accessed June 1, 2020.
16. Kong WH, Li Y, Peng MW, et al. SARS-CoV-2 detection in patients with influenza-like illness. *Nat Microbiol*. 2020;5(5):675-678. doi:10.1038/s41564-020-0713-1
17. Xu B, Gutierrez B, Mekaru S, et al. Epidemiological data from the COVID-19 outbreak, real-time case information. *Sci Data*. 2020;7(1):1-6. doi:10.1038/s41597-020-0448-0
18. Pan L, Mu M, Yang P, et al. Clinical Characteristics of COVID-19 Patients With Digestive Symptoms in Hubei, China. *Am J Gastroenterol*. 2020;115(5):766-773. doi:10.14309/ajg.0000000000000620
19. To KKW, Tsang OTY, Leung WS, et al. Temporal profiles of viral load in posterior

oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis*. 2020;20(5):565-574. doi:10.1016/S1473-3099(20)30196-1

20. Brownstein JS, Kleinman KP, Mandl KD. Identifying pediatric age groups for influenza vaccination using a real-time regional surveillance system. *Am J Epidemiol*. 2005;162(7):686-693. doi:10.1093/aje/kwi257
21. Han C, Duan C, Zhang S, et al. Digestive Symptoms in COVID-19 Patients With Mild Disease Severity. *Am J Gastroenterol*. April 2020:1. doi:10.14309/ajg.0000000000000664
22. 宣传处. *Hubei Government Situation Report, May 10, 2020*.; 2020. http://wjw.hubei.gov.cn/fbjd/tzgg/202005/t20200511_2266405.shtml.
23. 武汉新增5例确诊患者 来自同一小区-新华网. http://www.xinhuanet.com/politics/2020-05/11/c_1125968733.htm. Accessed June 1, 2020.
24. Butler D. When Google got flu wrong. *Nature*. 2013;494(7436):155-156. doi:10.1038/494155a
25. Santillana M, Zhang DW, Althouse BM, Ayers JW. What can digital disease detection learn from Google flu trends? *Am J Prev Med*. 2014;47(3):341-347. doi:10.1016/j.amepre.2014.05.020
26. 武汉：全市网民770万 超98%用手机上网_全国. https://www.sohu.com/a/284177554_119038. Accessed June 1, 2020.
27. Nagar R, Yuan Q, Freifeld CC, et al. A case study of the New York City 2012-2013 influenza season with daily geocoded Twitter data from temporal and spatiotemporal perspectives. *J Med Internet Res*. 2014;16(10):e236. doi:10.2196/jmir.3416
28. Lu T, Reis BY. Internet Search Patterns Reveal Clinical Course of Disease Progression for COVID-19 and Predict Pandemic Spread in 32 Countries. *medRxiv*. 2020:2020.05.01.20087858. doi:10.1101/2020.05.01.20087858

Figures

Figure 1: Parking lot volume in a Wuhan, China hospital and control site

Satellite images and counts of cars (red) and trucks (yellow) in 2 Wuhan, China, parking areas. Top panel is Wuhan Tianyou Hospital pre-epidemic (***A***), October 2019 (***B***) and during the height of the COVID-19 outbreak (***C***). Lower panel is the Huanan Seafood Market, a high-volume control site, in September 2019 (***D***) and February 2020 (***E***).

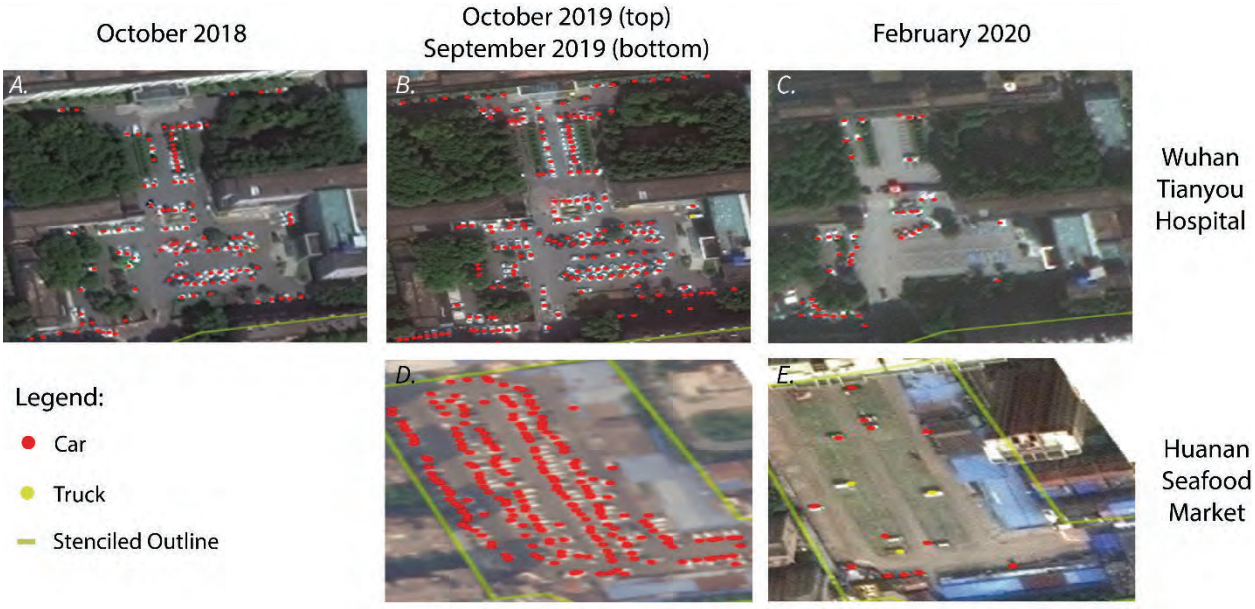
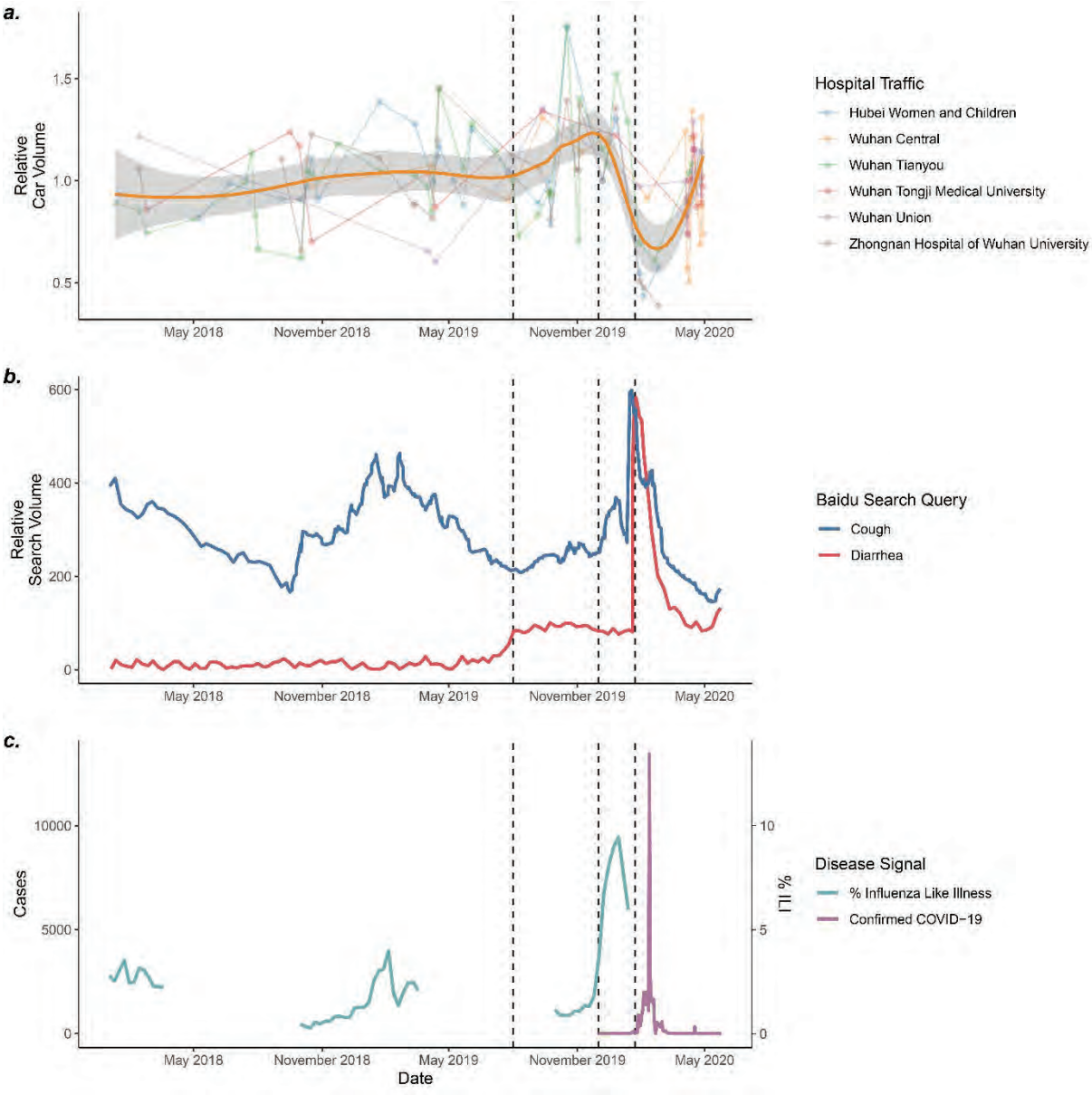


Figure 2: Time-series of different ILI, symptoms and surveillance signals, January 2018 - May 2020

Daily parking lot volume normalized to the same hospital's baseline (weekday, Saturday or Sunday) for 6 Wuhan, China Hospitals (***a***) from January 2018 - April 2020 and a fitted loess regression (***a***, orange) with standard errors. Relative Baidu search query volume is shown for the terms "cough" (***b***, blue) and "diarrhea" (***b***, red) between January 2018 - May 2020. Three yearly influenza-like-illness cycles (***c***, turquoise, right side axis) from two Wuhan hospitals plotted alongside confirmed COVID-19 case counts (***c***, purple, left side axis) extracted from open-source disease data. Date annotations (vertical dashed lines) on all three panels representing August 1, 2019 (rise in traffic and diarrhea signal), December 1, 2019 (date of first confirmed COVID-19 case) and January 23, 2020 (implementation of the Wuhan public health lockdown).



Exhibit

39

Table 1. Diarrhea Assessments Based on Patient Interviews and Documentation in Medical Records for 100 Patients Tested for *Clostridioides difficile* Infection

Variable	Total (N=100), No. (%)	Meeting Criteria for Diarrhea[a] (N=60), No. (%)	Not Meeting Criteria for Diarrhea[a] (N=40), No. (%)
Patient interview			
Self-reported diarrhea[b]	86 (86)	60 (100)	26 (65)
≥3 unformed stools per day	60 (60)[c]	60 (100)	0 (0)
<3 unformed stools per day	20 (20)	0 (0)	20 (50)
Formed stool but with increased frequency	6 (6)	0 (0)	6 (15)
No diarrhea (normal stool frequency and consistency)	14 (14)	0 (0)	14 (35)
Practitioner documentation[d]			
Diarrhea	75 (75)	51 (85)	24 (60)
No. of bowel movements	46 (46)	33 (55)	13 (33)
Consistency of stools	46 (46)	32 (53)	14 (35)
Nursing documentation			
Diarrhea	18 (18)	11 (18)	7 (18)
No. of bowel movements	20 (20)	11 (18)	9 (23)
Consistency of stools	24 (24)	12 (20)	12 (30)

[a]Diarrhea defined as 3 or more unformed (Bristol scale 6 or 7) stools in a 24-hour period as determined by patient interview.[b]Patients were first asked if they had diarrhea without any comment on how diarrhea should be defined.[c]Of 60 patients with ≥3 unformed stools per day, 8 (13%) did not meet criteria for clinically significant diarrhea because they had a clear alternative explanation for diarrhea (eg, laxatives, chronic diarrhea due to chronic pancreatitis).[d]Practitioners included physicians, nurse practitioners, and physician assistants.

stools. Education of patients could empower them to participate in efforts to reduce inappropriate CDI testing.[8,9] Education of personnel to obtain information on frequency and consistency of stools could improve the accuracy of diarrhea documentation.

Our study had some limitations. Only one healthcare facility was included. At the time of the study, no interventions were in place to limit inappropriate CDI testing. The lack of documentation by nurses in outpatient settings is not unexpected. Finally, in some cases, the information on bowel movements provided by patients or family members may have been inaccurate.

In conclusion, education of personnel and patients about the definition of clinically significant diarrhea and efforts to improve documentation of diarrhea are needed to support CDI diagnostic stewardship interventions.

Acknowledgments. We thank the patients who participated in the study.

Financial support. This work was supported by the Department of Veterans' Affairs.

Competing interests. C.J.D. has received research grants from Clorox, Pfizer, and PDI. All other authors report no potential conflicts relevant to this article.

References

1. McDonald LC, Gerding DN, Johnson S, *et al.* Clinical practice guidelines for *Clostridium difficile* infection in adults and children: 2017 update by the Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA). *Clin Infect Dis* 2018;66:987–994.
2. Kociolek LK, Gerding DN, Carrico R, *et al.* Strategies to prevent *Clostridioides difficile* infections in acute-care hospitals: 2022 update. *Infect Control Hosp Epidemiol* 2023;44:527–549.
3. Donskey CJ, Kundrapu S, Deshpande A. Colonization versus carriage of *Clostridium difficile*. *Infect Dis Clin N Am* 2015;29:13–28.
4. Curry SR, Hecker MT, O'Hagan J, *et al.* Natural history of *Clostridioides difficile* colonization and infection following new acquisition of carriage in healthcare settings: a prospective cohort study. *Clin Infect Dis* 2023. doi: 10.1093/cid/ciad142.
5. Kara A, Tahir M, Snyderman W, Brinkman A, Fadel W, Dbeibo L. Why do clinicians order inappropriate *Clostridium difficile* testing? An exploratory study. *Am J Infect Control* 2019;47:285–289.
6. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. *Scand J Gastroenterol* 1997;32:920–924.
7. DeBenedictus CM, Hecker MT, Zuccaro PD, John JP, Donskey CJ, Patel PK. What is the current state of patient education after *Clostridioides difficile* infection? *Infect Control Hosp Epidemiol* 2020;41:1338–1340.
8. Hecker MT, Son AH, Zuccaro P, Conti J, Donskey CJ. Real-world evaluation of a two-step testing algorithm for *Clostridioides difficile* infection. *Infect Control Hosp Epidemiol* 2023. doi: 10.1017/ice.2022.313.
9. Donskey CJ. Update on *Clostridioides difficile* infection in older adults. *Infect Dis Clin North Am* 2023;37:87–102.

Cost of personal protective equipment during the first wave of the coronavirus disease 2019 (COVID-19) pandemic

Ria Patel BS[1], Mirza Ali CIC[2], Susan C. Bleasdale MD[3] and Alfredo J. Mena Lora MD[2,3]

[1]Ross University School of Medicine, Miramar, Florida, [2]Saint Anthony Hospital, Chicago, Illinois and [3]University of Illinois at Chicago, Chicago, Illinois

To the Editor—As the world prepared and responded to the coronavirus disease 2019 (COVID-19) pandemic in early 2020, a rapid increase in demand for personal protective equipment (PPE) led to severe shortages worldwide. The PPE demand rose as a result of panic purchasing, hoarding, and misinterpretation of public health information.[1–3] This led to shortages so wide that the World Health Organization released several memorandums regarding 'rational use' of PPE to try and reconcile the spike in utlization of PPE as

Corresponding author: Alfredo J. Mena Lora; Email: amenalor@uic.edu

Cite this article: Patel R, *et al.* (2023). Cost of personal protective equipment during the first wave of the coronavirus disease 2019 (COVID-19) pandemic. *Infection Control & Hospital Epidemiology*, 44: 1897–1899, https://doi.org/10.1017/ice.2023.115

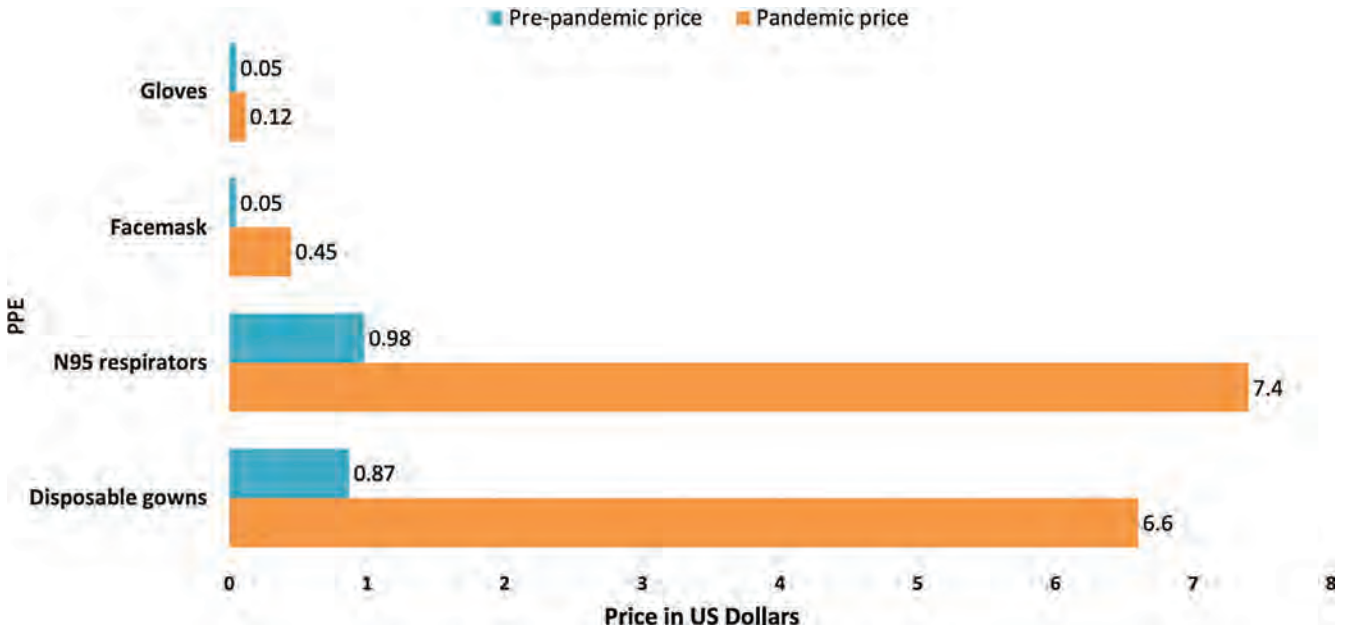


Fig. 1. Cost of personal protective equipment per unit in US dollars.

well as the inevitable increase in prices.[1,4] Our 151-bed safety-net hospital, like many other facilities, navigated a major rise in prices as part of our pandemic preparedness and response. In this study, we performed a review of PPE market prices paid by our facility during the first surge of the pandemic (April–June 2020). Our hospital is located in Chicago, Illinois. We have reported cost and emphasized the importance of PPE supply resilience for pandemic preparedness. The maximum cost per unit (CPU) of PPE was tabulated and compared with prepandemic (April–June 2019) prices to establish an increase in CPU in US dollars for the various PPE items. The maximum CPU of PPE was tabulated for each week and the average cost throughout the first wave of COVID-19 pandemic was calculated. Disposable gowns, washable gowns, N95 respirators, face masks, and gloves were included in our analysis.

PPE prices were significantly higher during the first wave of the pandemic compared to prepandemic prices (Fig. 1). The CPU for gloves increased from $0.05 to $0.12; the CPU for face masks increased from $0.05 to $0.45; the CPU for N95 respirators increased from $0.98 to $7.4; and the CPU for disposable gowns increased from $0.87 to $6.60. The CPU for gloves averaged 2.5 times higher than prepandemic prices. The CPU for face masks peaked at $0.55, 11 times higher and an average of 9 times higher than the prepandemic price. N95 respirators had a peak CPU of $12, and the average CPU was 8 times higher than prepandemic prices. The average CPU for gowns was 7.5 times higher than prepandemic prices; the CPU for disposable gowns peaked at $12 during the first week of March 2020, a price 13.7 times higher than the prepandemic price.

Before the pandemic, hospitals were spending ~$7 per patient on PPE, with an increase to $20.40 during the first wave of COVID-19 in the spring of 2020.[5] The Society for Healthcare Organization Procurement Professionals conducted a study during the same period showing that PPE costs increased 1,064% in their 5,000 skilled nursing facilities and assisted living centers across the United States.[2] Multiple factors likely contributed to high prices, including demand shock, disrupted supply chains, and a rush to acquisition by healthcare systems and the general population alike. The global PPE supply chain did not properly operate to meet the demands of healthcare systems across the world during the first wave of the COVID-19 pandemic. In 2020, many factors such as the shortage of raw materials, export bans, and restraints in logistics contributed to 4–6 month backlogs for global supply orders of PPE.[6]

Several major sources of PPE backlogs were identified by the Asian Development Bank during the early wave of the pandemic in 2020. N95 mask surges led to shortages in nonwoven polypropylene. Export bans in 22 countries limited PPE access to millions of people worldwide. Transport and shipping constraints were implemented. Fewer workplace personnel were available due to COVID-19. Availability of various shipping containers was lower, and more.[6] Along with price gouging and competitive bidding among hospitals, these factors inevitably contributed to soaring PPE prices. Cohen et al[3] elaborated on 4 other contributing factors to PPE shortages: the way that hospitals budget for PPE, domestic demand shocks, federal government failures, and disruptions to global supply chain.[3]

The PPE supply challenges during COVID-19 led to several important conversations regarding policy changes and mitigation strategies. The *International Journal of Operations and Production Management* identified 4 specific strategies for pandemic-specific supply-chain management: global PPE standards, production changeover, joint procurement, and multiple sourcing.[7] The Asian Development Bank also outlined several policy recommendations to ensure PPE preparedness for a global pandemic: monitoring PPE use and visibility of orders, improving supply systems and sharing responsibility, improving domestic manufacturing surge capacity during the event, strengthening trade finance programs for micro, small, and medium-sized enterprises, and more.[6] To develop resilient supply chains, local capacity for essential items should be developed, along with an increase and diversification in international production sites.[8]

The COVID-19 pandemic has highlighted the need for policy reconciliation to prevent future supply chain failures and inevitably price surging for essential materials such as PPE.

The Society for Healthcare Epidemiology of America (SHEA) Outbreak Response and Incident Management guidelines provide a high-level overview of incident management for infectious diseases outbreaks.[9] Our facility navigated these challenges by centralizing and

https://doi.org/10.1017/ice.2023.115 Published online by Cambridge University Press

securing our PPE stocks, ensuring accountability and evidence-based use of our supplies, and diversifying our procurement as much as possible to ensure adequate stocks. Resilient supply chains and standardized guidelines for PPE reuse may be necessary for future pandemics.

This study had several limitations. For example, the data presented were limited to the geographic scope of Chicago, Illinois. Despite these limitations, the results of this study are generalizable to hospitals across the United States.

Acknowledgments.

Financial support. No financial support was provided relevant to this article.

Competing interests. All authors report no conflicts of interest relevant to this article.

References

1. Battista RA, Ferraro M, Piccioni LO, Malzaneni GE, Bussi M. Personal protective equipment (PPE) in COVID 19 pandemic: related symptoms and adverse reactions in healthcare workers and general population. *J Occupat Environ Med* 2021;63:e80–e85.
2. Berklan JM. Analysis: PPE costs increase over 1,000% during COVID-19 crisis. McKnight's Long-Term Care News website. https://www.mcknights.com/news/analysis-ppe-costs-increase-over-1000-during-covid-19-crisis/. Published April 9, 2020. Accessed January 6, 2023.
3. Cohen J, van der Meulen Rodgers Y. Contributing factors to personal protective equipment shortages during the COVID-19 pandemic. *Prev Med* 2020;141:106263.
4. Rational use of personal protective equipment (PPE) for coronavirus disease (COVID-19): interim guidance, 19 March 2020. World Health Organization website. https://apps.who.int/iris/handle/10665/331498. Published January 1, 1970. Accessed January 6, 2023.
5. The current state of PPE costs: Are providers out of the woods? Premier website. https://premierinc.com/newsroom/blog/the-current-state-of-ppe-costs-are-providers-out-of-the-woods-1. Published January 6, 2023. Accessed January 6, 2023.
6. Park C-Y, Kim K, Roth S. Global shortage of personal protective equipment amid COVID-19: supply chains, bottlenecks, and policy implications. Asian Development Bank website. https://www.adb.org/publications/shortage-ppe-covid-19-supply-chains-bottlenecks-policy. Published July 28, 2020. Accessed January 6, 2023.
7. Sigala IF, Sirenko M, Comes T, Kovács G. Mitigating personal protective equipment (PPE) supply chain disruptions in pandemics—a system dynamics approach. *Int J Operat Prod Manag* 2022. doi: 10.1108/IJOPM-09-2021-0608.
8. Gereffi G. What does the covid-19 pandemic teach us about global value chains? The case of medical supplies. *J Int Bus Pol* 2020;3:287–301.
9. Banach DB, Johnston BL, Al-Zubeidi D, *et al.* Outbreak response and incident management: SHEA guidance and resources for healthcare epidemiologists in United States acute-care hospitals. *Infect Control Hosp Epidemiol* 2017;38:1393–1419.

https://doi.org/10.1017/ice.2023.115 Published online by Cambridge University Press

Exhibit

40

Asymptomatic SARS-CoV-2 infection: A systematic review and meta-analysis

Pratha Sah[a], Meagan C. Fitzpatrick[a,b], Charlotte F. Zimmer[a], Elaheh Abdollahi[c], Lyndon Juden-Kelly[c], Seyed M. Moghadas[c], Burton H. Singer[d,1], and Alison P. Galvani[a]

[a]Center for Infectious Disease Modeling and Analysis, Yale School of Public Health, New Haven, CT 06520; [b]Center for Vaccine Development and Global Health, University of Maryland School of Medicine, Baltimore, MD 21201; [c]Agent-Based Modelling Laboratory, York University, Toronto, ON M3J 1P3, Canada; and [d]Emerging Pathogens Institute, University of Florida, Gainesville, FL 32610

Contributed by Burton H. Singer, July 8, 2021 (sent for review May 19, 2021; reviewed by David Fisman and Claudio Jose Struchiner)

Quantification of asymptomatic infections is fundamental for effective public health responses to the COVID-19 pandemic. Discrepancies regarding the extent of asymptomaticity have arisen from inconsistent terminology as well as conflation of index and secondary cases which biases toward lower asymptomaticity. We searched PubMed, Embase, Web of Science, and World Health Organization Global Research Database on COVID-19 between January 1, 2020 and April 2, 2021 to identify studies that reported silent infections at the time of testing, whether presymptomatic or asymptomatic. Index cases were removed to minimize representational bias that would result in overestimation of symptomaticity. By analyzing over 350 studies, we estimate that the percentage of infections that never developed clinical symptoms, and thus were truly asymptomatic, was 35.1% (95% CI: 30.7 to 39.9%). At the time of testing, 42.8% (95% prediction interval: 5.2 to 91.1%) of cases exhibited no symptoms, a group comprising both asymptomatic and presymptomatic infections. Asymptomaticity was significantly lower among the elderly, at 19.7% (95% CI: 12.7 to 29.4%) compared with children at 46.7% (95% CI: 32.0 to 62.0%). We also found that cases with comorbidities had significantly lower asymptomaticity compared to cases with no underlying medical conditions. Without proactive policies to detect asymptomatic infections, such as rapid contact tracing, prolonged efforts for pandemic control may be needed even in the presence of vaccination.

asymptomatic fraction | presymptomatic | silent transmission | novel coronavirus | comorbidity

COVID-19 surveillance provides real-time information about the epidemiological trajectory of the pandemic, informing risk assessments and mitigation policies around the world. Given that COVID-19 surveillance systems predominantly rely on symptom-based screening, the prevalence of asymptomatic infection is often not fully captured. Cross-sectional surveys, such as mass testing once an outbreak is identified, do not distinguish the truly asymptomatic from the presymptomatic. Often, the follow-up period after testing is too brief to ascertain whether patients subsequently develop symptoms. The percentage of silent infections identified by such studies is thus context specific, as it depends on the setting, phase of the epidemic, and efficiency of contact tracing. By contrast, the prevalence of truly asymptomatic infections should be stable across similar demographic settings, regardless of epidemiological trajectory and contact tracing.

Compounded by ambiguities about the different clinical manifestations of the disease, which can lead to misinterpretation of clinical and epidemiological studies (1), there have been substantial aberrations in reports and media coverage claiming the asymptomatic percentage to be as low as 4% (2, 3) or as high as 80 to 90% (4, 5). Similarly, the US Centers for Disease Control and Prevention guidelines for COVID-19 pandemic forecasting offer wide bounds for the asymptomatic percentage, ranging from 10 to 70% (6).

Previous meta-analyses of 41 studies (7), 13 studies (8), and 79 studies (9) estimate pooled asymptomaticity ranging from 16 to 20%. Two methodological issues limit the accuracy of these studies. First, pooled asymptomaticity reported in these studies is likely biased downward because they did not account for study designs which have a higher representation of cases experiencing symptoms (10). Second, one of the meta-analyses (7) did not consider biases in reported asymptomaticity that can arise from inadequate longitudinal follow-up. Studies that assess the symptom profile only at the time of testing or do not follow up symptoms for a sufficiently long time period cannot distinguish presymptomatic from asymptomatic infection, overestimating those that are truly asymptomatic.

Accurate estimates of true disease prevalence, including asymptomatic infections, are essential to calculate key clinical parameters, project epidemiological trajectories, and optimize mitigation measures. Clinical evidence indicates that viral loads among asymptomatic and symptomatic infections may be comparable (11–15). Unaware of their risk to others, individuals with silent infections are likely to continue usual behavior patterns. Accounting for silent severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections in the assessment of disease control measures is necessary to interrupt community transmission (16). Although the discrepancy between reported incidence and seroprevalence gives a sense of the extent of asymptomaticity, not

Significance

Asymptomatic infections have been widely reported for COVID-19. However, many studies do not distinguish between the presymptomatic stage and truly asymptomatic infections. We conducted a systematic review and meta-analysis of COVID-19 literature reporting laboratory-confirmed infections to determine the burden of asymptomatic infections and removed index cases from our calculations to avoid conflation. By analyzing over 350 papers, we estimated that more than one-third of infections are truly asymptomatic. We found evidence of greater asymptomaticity in children compared with the elderly, and lower asymptomaticity among cases with comorbidities compared to cases with no underlying medical conditions. Greater asymptomaticity at younger ages suggests that heightened vigilance is needed among these individuals, to prevent spillover into the broader community.

Author contributions: P.S., M.C.F., S.M.M., and A.P.G. designed research; P.S., M.C.F., C.F.Z., E.A., L.J.-K., S.M.M., B.H.S., and A.P.G. performed research; P.S., M.C.F., C.F.Z., E.A., L.J.-K., S.M.M., and A.P.G. analyzed data; and P.S., M.C.F., S.M.M., B.H.S., and A.P.G. wrote the paper.

Reviewers: D.F., University of Toronto; and C.J.S., Fundacao Getulio Vargas.

The authors declare no competing interest.

See online for related content such as Commentaries.

[1]To whom correspondence may be addressed. Email: bhsinger@epi.ufl.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2109229118/-/DCSupplemental.

Published August 10, 2021.

POPULATION BIOLOGY

all symptomatic cases are reported, and not all asymptomatic cases (for instance, those identified on the basis of exposure) are missed. Consequently, it is not sufficient to simply compare the reported cases to results from seroprevalence studies. We therefore conducted a systematic review and meta-analysis of COVID-19 literature reporting laboratory-confirmed infections to estimate the percentage of SARS-CoV-2 infections that are truly asymptomatic. We also investigated differences in asymptomaticity with respect to age, sex, comorbidity, study design, publication date, duration of symptom follow-up, geographic location, and setting.

Results

We identified a total of 114,124 abstracts based on our search criteria. After excluding duplicate and irrelevant studies, we used 390 in our meta-analyses (Fig. 1 and *SI Appendix*, Table S2). Most studies were conducted in China (n = 104, 27%), followed by the United States (n = 74, 19%), Italy (n = 21, 5%), and South Korea (n = 13, 3%). These studies included a total of 104,058 laboratory-confirmed COVID-19 cases, of which 25,050 exhibited no symptoms at the time of testing and 7,220 remained asymptomatic. We identified 170 studies that reported asymptomatic infections (11–13, 17–183), 332 studies that reported silent infections at the time of testing (10–12, 14, 17–20, 23–27, 31, 32, 35–40, 42–44, 46, 47, 49, 50, 52, 53, 56–58, 60–66, 68, 69, 73–75, 77–79, 81, 84, 87, 90–94, 97, 99, 101, 103, 104, 106, 111, 113–116, 118, 119, 121–123, 125, 127, 128, 131, 133, 135, 137, 138, 140, 143, 145, 146, 148–152, 154, 156, 158, 160–163, 166–170, 172–174, 176, 177, 179, 180, 182–405), and 143 that delineated presymptomatic and asymptomatic infections by following-up with those silently infected (11–13, 17–20, 22–29, 31–33, 35–40, 42–44, 46–54, 56–70, 72–75, 77–81, 83, 84, 87, 89–94, 96, 97, 99, 101, 103, 104, 106–109, 111–119, 121–125, 127–129, 131, 133, 135–138, 140, 141, 143, 145–156, 158–164, 166–170, 172–174, 176–183). Among the studies that reported follow-up of clinical symptoms after testing, 11.0% reported at time points at 1 wk to 2 wk, 33.8% reported at 2 wk to 3 wk, and 55.2% reported longer than 3 wk. Among the studies that reported asymptomatic infections, 58.8% reported zero index cases, either because cases were identified through a screening design or because the study only reported the cases that were identified through contact tracing. Of the 41.2% studies that reported data on index cases, these included household members, long-term care residents, members of the community, or travelers returning from COVID-19 hotspots (*SI Appendix*, Table S1).

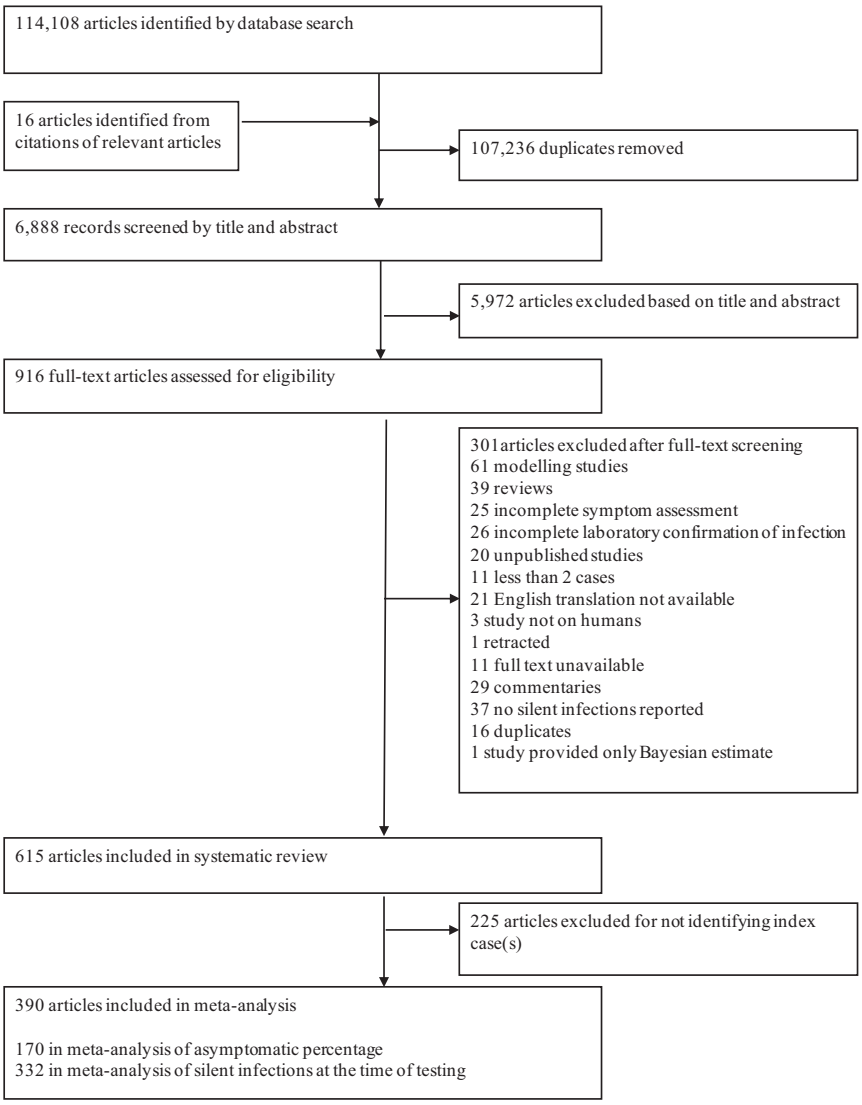


Fig. 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram showing the numbers of studies screened and included in the meta-analysis.

A summary of the risk of bias assessment is presented in *SI Appendix*, Table S2. Out of the 170 studies included in the calculation of asymptomaticity, 75 had low risk of bias, 10 had moderate risk of bias, and 85 had serious risk of bias.

The percentage of cases that were truly asymptomatic among laboratory-confirmed cases was 35.1% (95% CI: 30.7 to 39.9%; Fig. 2). By contrast, a larger percentage of cases exhibited no symptoms at the time of testing (42.8%, 95% prediction interval: 5.2 to 91.1%) due to mischaracterization of presymptomatic cases as asymptomatic. To investigate the degree of mischaracterization, we considered a subset of studies that reported symptoms both at the time of testing and a minimum of 7 d after. Within this subset of studies, 31.8% (95% prediction interval: 5.6 to 78.7%) of cases exhibiting no symptoms at the time of testing progressed to develop symptoms. The percentage of truly asymptomatic cases among these studies was therefore 36.9% (95% CI: 31.8 to 42.4%), similar to that estimated for all studies reporting asymptomatic infections.

These estimates were obtained after removing index cases from our calculations, correcting bias toward overrepresentation of symptomatic cases that would lead to underestimation of asymptomaticity. Without excluding index cases, estimates of asymptomatic infections using our two complementary approaches would be 27.8% (95% CI: 24.3 to 31.7%) and 29.4 (95% CI: 25.2 to 33.9%). To evaluate the impact of sample selection bias arising from higher participation among those experiencing symptoms, we next restricted our analysis to 25 studies in which complete screening of every individual present at the setting was performed. The pooled asymptomaticity among this smaller subset of studies was 47.3% (95% CI: 34.0 to 61.0%).

We found a statistically significant trend toward a lower asymptomatic percentage with increasing age ($P < 0.01$; Table 1). In pairwise comparisons, the asymptomatic percentage was significantly lower for the elderly, at 19.7% (95% CI: 12.7 to 29.4%) compared with 46.7% (95% CI: 32.0 to 62.0%) for children ($P < 0.01$). Asymptomaticity also varied across study settings ($P = 0.03$; Table 1). In particular, studies on long-term care facilities reported lower asymptomaticity compared with studies on healthcare facilities ($P = 0.04$) and household transmission ($P = 0.04$). We found no association between asymptomatic percentage and geographic location, study design, follow-up duration, or publication date (Table 1). We found that asymptomaticity in males was similar to that in females (log incidence rate ratio [IRR] 0.09, 95% CI −0.07 to 0.25, $P = 0.27$; *SI Appendix*, Fig. S1). Cases with comorbidities had lower asymptomaticity compared to cases with no underlying medical conditions (log IRR −0.43, 95% CI −0.82 to −0.04, $P = 0.03$; *SI Appendix*, Fig. S2).

Egger's test for asymptomatic percentage was significant ($P = 0.04$; *SI Appendix*, Fig. S3), providing evidence of potential small-study effects. We therefore conducted a sensitivity analysis by excluding studies with relatively small sample sizes (less than 10 infections). The pooled estimate in the restricted meta-analysis (33.1%; 95% CI: 28.0 to 38.5%) was similar to our original estimate, suggesting that our estimates are robust to publication bias.

Discussion

The SARS-CoV-2 pandemic infected more than 80 million people within a year and is still spreading rapidly despite widespread control efforts. The elements of the global response are similar to those deployed during the SARS-CoV-1 outbreak: detecting new cases through symptom-based surveillance, subsequent testing, and isolation of confirmed cases. In 2002, these measures achieved containment within 8 mo and fewer than 8,500 cases worldwide. Given that the aerosol and surface stability of the two viruses are similar (406), a crucial difference between the two outbreaks could be the role of silent infections in propagating transmission chains. Multiple clinical studies have indicated that viral loads in asymptomatic and symptomatic

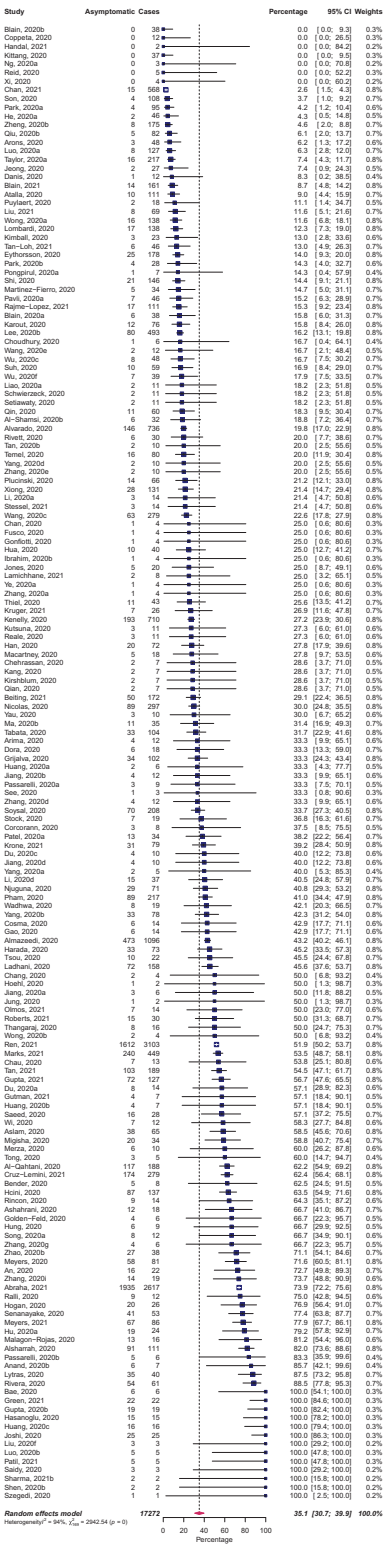


Fig. 2. Pooled percentage of laboratory-confirmed COVID-19 cases which remained asymptomatic. Studies that did not report follow-up of silent infections or failed to identify index cases were excluded from the analysis.

infections of COVID-19 may be similar (11–14, 354). Furthermore, the presymptomatic phase of SARS-CoV-2 is highly infectious (53), and transmission from those in this phase may be responsible for more than 50% of incidence (16). This is a

Table 1. Pooled estimates for percentages of all positive cases which remain asymptomatic stratified by age, gender, publication date, symptom follow-up duration, study design, and study setting

	n	Estimate (%)	CI (95%)	*P* value (test of overall effect)
Age class				**<0.01**
Children (0 y to 18 y)	18	46.7	32.0 to 62.0	
Adults (19 y to 59 y)	17	32.1	22.2 to 43.9	
Elderly (≥60 y)	17	19.7	12.7 to 29.4	
Study design				0.10
Population screening	102	38.2	32.0 to 44.8	
Others	68	30.7	24.8 to 37.4	
Publication date				0.18
January–April 2020	27	34.8	23.6 to 47.9	
May–August 2020	69	29.5	24.2 to 35.4	
September–December 2020	50	41.1	31.4 to 51.4	
January–April 2021	24	38.4	25.6 to 53.1	
Symptom follow-up duration				0.07
7 d to 21 d	73	40.6	32.9 to 48.6	
21+ d	90	32.1	27.0 to 37.7	
Setting				**0.03**
Community	39	34.0	25.3 to 43.8	
Healthcare facility	81	38.5	31.6 to 45.9	
Household	18	42.5	30.9 to 54.9	
Long-term care facility	15	17.8	9.7 to 30.3	
Others	17	38.4	23.5 to 55.9	
Geographic location				0.78
China	50	33.6	26.1 to 42.0	
United States	28	33.3	22.6 to 46.1	
Others	92	36.8	30.4 to 43.6	

Stratifications with statistically significant subgroup differences ($P < 0.05$) are in bold.

striking difference from SARS-CoV-1 in which the infectiousness peaked at 12 d to 14 d after symptom onset (407). Although silent infections of SARS-CoV-1 were reported, no known transmission occurred from silently infected or even mildly symptomatic SARS cases.

Since the emergence of COVID-19, there has been much speculation about the silent transmission of the disease. Cross-sectional studies testing exposed individuals who do not exhibit symptoms often conflate asymptomatic infections with those in the presymptomatic phase, leading to substantial overestimation of asymptomatic infection. Longitudinal studies without sufficient follow-up similarly lead to overestimation of asymptomaticity (408). Additionally, inconsistent use of terminology has led to confusion, particularly when distinguishing infections which are silent at the time of testing from those which are truly asymptomatic (4, 5). A previous meta-analysis, for example, incorrectly includes infections in the presymptomatic phase to calculate pooled estimate of asymptomatic percentage (409). By contrast, several studies conducted early in the pandemic reported few asymptomatic infections, primarily due to restrictive testing criteria which focused on testing of severe cases that required hospitalization (410, 411). Inaccuracy in either direction is detrimental for public health. Overestimation of asymptomaticity engenders a perception that SARS-CoV-2 is less virulent, whereas underestimation skews key epidemiological parameters such as infection fatality rate and hospitalization rate upward, leading to suboptimal policy decisions.

To robustly estimate the asymptomatic percentage from studies with varying degrees of methodological vigor, we conducted two separate meta-analyses. In the first analysis, we estimated the asymptomatic percentage as 35.1% (95% CI: 30.7 to 39.9%), by including all studies with a duration of follow-up sufficient to identify asymptomatic infections. In the second analysis, we only included studies that both delineated silent infections at the time of testing and conducted follow-up to distinguish the presymptomatic stage from asymptomatic infections. With this analysis, we estimated the asymptomatic percentage as 36.9% (95% CI: 31.8 to 42.4%). Our estimates have overlapping CIs, which suggests that our pooled analysis is robust to methodological differences in symptom assessment. Our estimates are higher than the 15.6% (95% CI: 10.1 to 23.0%), 17% (95% CI: 14 to 20%), and 20% (95% CI: 17 to 25%) reported by three previous meta-analyses using 41 studies (7), 13 studies (8), and 79 studies (9). In large part, this difference arises because we excluded index cases from our calculation, correcting a bias that leads to underestimation of asymptomaticity. Our estimates of asymptomatic percentage without excluding index cases were 27.8% and 29.4%, for our two approaches. The lower bounds of 24% and 25%, for the two analyses overlaps with the range of the previous largest meta-analysis. Compared with other respiratory infections, the lower bound of our analyses is higher than the 13 to 19% estimated for influenza (412, 413), and the 13% for SARS-CoV-1 (414).

We found that 42.8% (95% prediction interval: 5.2 to 91.1%) of infections were silent at the time of testing. These cases have been incorrectly referred to as asymptomatic in previous studies (4, 5, 189, 239). This rate is context specific, as it is likely influenced by the association between symptomaticity and the time window when an infection is detectable or tested by RT-PCR. Additionally, the proportion of silent infections at the time of testing is highly sensitive to the efficiency of contact tracing. If most contacts are identified and tested swiftly, then nearly all infections will be silent at the time of testing. By contrast, if contact tracing is slow and incomplete, then a larger fraction of individuals will have developed symptoms by the time they are approached for testing, and a smaller proportion of those tested will be symptom-free. Reports of silent infections at the time of testing are also likely impacted by epidemic trajectory largely due to the predominance of recent infections in samples taken during

the growth phase, in contrast with a higher proportion of older infections in samples taken during the declining phase. Unbiased measures of asymptomaticity, on the other hand, should be consistent across similar demographic settings, regardless of contact tracing and epidemic trajectory.

Several gaps remain in our understanding of asymptomatic carriage of COVID-19. Particularly, it is unclear why certain infections remain asymptomatic while the majority develop clinical symptoms. Our results indicate that children have greater asymptomaticity compared to the elderly. We also found that cases with comorbidities have lower asymptomaticity compared with cases with no underlying medical conditions. Additionally, studies on long-term care facilities reported lower asymptomaticity compared to other study settings. Given that the risk of severe illness is high among the elderly, the age association identified by our study implies that absence of symptoms may correlate with the tendency of developing milder symptoms. Case severity in SARS-CoV-2 patients has been linked to a cytokine storm which occurs more frequently in elderly patients (415, 416). Genetic (417), environmental risk factors, sex-linked differences (418), and cross-reactive immunity (419) might also contribute, although no studies have unequivocally demonstrated their association with either symptom status or severity.

Higher representation of asymptomatic SARS-CoV-2 infections among younger people has grave implications for control policies in daycares, schools, and universities. Settings with close, extensive contact among large groups of younger individuals are particularly susceptible to superspreader events of COVID-19 which may go undetected if surveillance focuses on symptomatic cases. This close congregation of relatively large groups similarly explains why influenza, mumps, and measles often spread more rapidly in schools and college campuses than in the broader community (420–422). As schools and universities convene in the midst of the COVID-19 pandemic, campus outbreaks are increasingly reported (423). Although COVID-19 severity is lower among young people, campus transmission with a large undetected component could more easily bridge to the rest of the population, fueling local and regional resurgence.

Our meta-analyses are subject to limitations, many related to the unprecedented pace of clinical research since the emergence of COVID-19. First, we found considerable heterogeneity in the percentage of asymptomatic infections. Subgroup analysis revealed that studies with longer follow-up reported lower asymptomaticity. Second, all reports of asymptomatic cases are confounded by the subjective and shifting definition of symptoms. For instance, the list of clinical manifestations associated with COVID-19 has expanded since the initial definitions (424). These changing definitions impact the classification of infections as asymptomatic or silent, and the more limited suite of symptoms initially considered indications of COVID-19 could bias early studies toward higher percentages in these categories. Nonetheless, we found no statistically significant differences in asymptomatic percentage when we stratified studies based on publication date. Third, in the studies included in our meta-analysis, it is possible that early mild symptoms occurring before a positive PCR test might go unrecorded, biasing the studies toward higher asymptomaticity. Fourth, although we corrected for the bias introduced by inclusion of predominantly symptomatic index cases, our estimates are still likely affected by sample selection bias, as participation is expected to be highest among those experiencing symptoms (10). Additionally, factors such as socioeconomic position, occupation, ethnicity, place of residence, internet and technological access, and scientific and medical interest could have contributed to nonrandom enrollment (425). To evaluate the effect of these biases, we calculated the pooled asymptomatic percentage using 25 studies that reported screening of all individuals in the study setting. Asymptomaticity among this smaller subset of studies was 47.3% (95% CI: 34.0 to 61.0%), with CIs that overlap with our primary analysis but the point estimate is higher than the base case CI. We therefore cannot rule out nonrandom sampling as a source of bias for estimation of the asymptomatic percentage.

In our meta-analysis, we excluded 225 studies that did not identify index cases. Additionally, 223 studies reported silent infections at the time of testing but were excluded from analysis of asymptomaticity for not reporting symptom assessment during follow-up for at least 7 d or for not specifying the duration of follow-up. Large-scale longitudinal surveys should prioritize the inclusion of these data to facilitate accurate estimation of the asymptomatic percentage. At minimum, such studies should report the number of index cases among their study participants, the clinical symptom status of individuals at the time of testing, the duration of symptom follow-up, and symptom status during the follow-up. Ideally, studies would additionally provide a full symptom profile both at time of testing and by the end of follow-up, to facilitate reclassification as case definitions are updated.

Estimating the extent of COVID-19 asymptomaticity is critical for calculating key epidemiological characteristics, quantifying the true prevalence of infection, and developing appropriate mitigation efforts. This meta-analysis also establishes a baseline for asymptomaticity, prior to widespread vaccination coverage. Amid concerns that vaccines may be less protective against infection than disease, widespread vaccination coverage may soon lead to a rise in the percentage of infections that present asymptomatically. The high prevalence of silent infections even at baseline, coupled with their transmission potential, necessitates accelerated contact tracing, testing, and isolation of infectious individuals, as symptom-based surveillance alone is inadequate for control.

POPULATION BIOLOGY

Methods

Definition of Silent, Asymptomatic, and Presymptomatic Infection. We defined silent infections as laboratory-confirmed COVID-19 cases that did not exhibit any clinical symptoms, including fever, upper respiratory symptoms, pneumonia, fatigue, headache, myalgia, dehydration, or gastrointestinal dysfunction, at the time of testing. Asymptomatic infections include those that continued to exhibit no clinical symptoms during at least 7 d of follow-up after testing. Presymptomatic cases were those that developed clinical symptoms subsequent to initial testing. The presymptomatic stage begins with the start of infectiousness and ends with the onset of symptoms (426).

Search Strategy and Selection Criteria. We conducted a systematic review to identify studies reporting laboratory-confirmed COVID-19 cases without symptoms at the time of testing. Our search was inclusive of all studies that provided data regarding cases that were asymptomatic, presymptomatic, or both. We finalized systematic search criteria on May 1, 2020, and study collection was initiated by searching PubMed, EMBASE, Web of Science, and the World Health Organization Global Research Database on COVID-19 (427) weekly from inception through April 2, 2021, with no language restrictions. Our search terms included "SARS-CoV-2," "novel coronavirus," "coronavirus 2019," "COVID-19," "COVID 2019" AND "asymptomatic," "no symptoms," "presymptomatic," "paucisymptomatic," "sub-clinical," "silent transmission," "silent infection," "without any symptoms," and "without symptoms" (*SI Appendix*, Table S1). All studies of any design that included these terms, were published after January 1, 2020, and described the symptom status of COVID-19 cases were considered in the screening step. No changes were made to the search criteria after the study initiation on May 1, 2020. The study protocol is available in the Open Science Framework online public database, registration DOI: 10.17605/OSF.IO/ZCJ62.

All articles were double-screened (by P.S. and C.F.Z.) based on the title and abstract. Studies were excluded if they were 1) duplicate publications, 2) editorials, reviews, discussions, or opinion pieces, 3) ambiguous about the presence of silent infection, 4) modeling studies without primary data, 5) based on fewer than two cases, 6) not conducted in humans, or 7) retracted. All identified full-text articles were reviewed by P.S. and C.F.Z. For each full-text article, we manually searched references for additional relevant studies. Studies included in our meta-analysis either reported laboratory confirmations of COVID-19 at a single time point, providing a snapshot of disease prevalence in the study subjects, or reported longitudinal data over a period of follow-up.

Risk of bias was assessed independently by two authors, and consensus was achieved through discussion. We adapted the ROBINS-I checklist (428) to include seven items: 1) enrollment of all patients satisfying the criteria for

inclusion, 2) enrollment of cases regardless of symptom status, 3) confirmation of cases using RT-PCR, 4) symptoms monitored by clinicians rather than self-reporting, 5) symptom assessment at the end of the follow-up period, 6) symptom follow-up duration of at least 7 d, and 7) loss to follow-up less than 5%.

Data Analysis. We conducted a meta-analysis using the studies identified through our systematic review to determine the prevalence of those truly asymptomatic among infected individuals. To delineate true asymptomaticity from the combination of asymptomatic and presymptomatic infections, we pursued two complementary analyses: 1) a single-step analysis based on reports of those who were asymptomatic at the end of a follow-up period and 2) a two-step analysis first evaluating the percentage of infections without symptoms at the time of testing and then assessing asymptomaticity by subtracting those that progressed to develop symptoms. In the single-step analysis, we calculated asymptomaticity as the percentage of confirmed COVID-19 cases that continued to exhibit no clinical symptoms for at least 7 d after testing, whether or not symptom status was reported specifically at the time of testing. In the two-step analysis, we focused on a subset of studies that distinguished asymptomatic cases from those that were presymptomatic by reporting symptoms at time of testing as well as conducting follow-up of symptoms for at least 7 d after testing. In both analyses, we removed index case(s) from the denominator of our calculations to minimize representational bias that would result in overestimation of symptomaticity. As a sensitivity analysis, we repeated our calculations including index cases. For studies that did not follow a population screening design, we assumed that single infections without an epidemiological link were necessarily detected due to their symptoms. Therefore, we subset the calculations to include only those infections which were part of a cluster.

To calculate pooled estimates, study outcomes were logit transformed, each study was assigned a weight using the inverse variance method (429), the DerSimonian–Laird estimator was applied to evaluate between-study variance (430), and the Clopper–Pearson method was used to determine CIs (431). Given heterogeneity in asymptomatic percentages estimated across studies, we used a random-effects meta-analysis model, applying the Hartung and Knapp (432) method to adjust test statistics and CIs for the random effect. We evaluated small-study effects visually with a contour-enhanced funnel plot and statistically with Egger's test (433). As a sensitivity analysis, we excluded studies with a small sample size (<10 infections), and we considered whether their removal impacted the pooling of results.

We conducted subgroup analysis stratified by age class, study design (population screening or not), publication date, duration of symptom follow-up, geographic location, and setting (community, healthcare facility, household, long-term care facilities, and other which encompassed schools, ships, conference, call centers, labor and delivery units, homeless shelters, and detention facilities). For subgroup analysis involving age class, we selected studies where all confirmed cases were either children (0 y to 18 y), adults (19 y to 59 y) or the elderly (≥60 y). We evaluated sex-based differences in asymptomaticity by selecting only those studies that stratified asymptomatic cases with respect to sex. For each of these studies we calculated the IRR, which was the ratio of the asymptomatic percentage in males relative to that in females. A similar analysis was performed to evaluate the asymptomaticity in cases with comorbidity relative to those without.

We next evaluated the impact of sample selection bias arising from higher participation among those experiencing symptoms in studies with voluntary participation. In this analysis, we calculated the pooled asymptomaticity after restricting to a smaller subset of studies that performed screening of every individual at the study setting. To avoid age-dependent bias in asymptomaticity, we removed studies where all participants belonged to a single age class (children, adults, or the elderly). Out of the 25 studies selected, 7 studies performed screening of all close household contacts (64, 80, 83, 103, 117, 131), 3 screened all flight passengers (28, 84, 91), and 2 screened all members of a tourist/pilgrim group (94, 129). Others were based on screening of healthcare workers (25, 110), inpatients admitted for non–COVID-19 reasons (19, 50, 59, 72, 108, 113), rigorously community screening (82, 166), travelers (18, 180), and those associated with a detention facility (92).

The meta-analysis and subgroup analyses were conducted using the metaprop function from the R package meta. Meta-analyses of sex-based and comorbidity-based differences in asymptomaticity were performed using the rma function from the R package metafor.

Data Availability. All study data are included in the article and *SI Appendix*.

ACKNOWLEDGMENTS. A.P.G. acknowledges funding from NSF Expeditions Grant 1918784, NIH Grant 1R01AI151176-01, NSF Grant RAPID-2027755, and the Notsew Orm Sands Foundation. S.M.M. was supported by the Canadian Institutes of Health Research [OV4 – 170643, COVID-19 Rapid Research] and Natural Sciences and Engineering Research Council of Canada Emerging Infectious Diseases Modelling Initiative (NSERC EIDM), Mathematics for Public Health (MfPH) grant.

1. X. Hao *et al*., Reconstruction of the full transmission dynamics of COVID-19 in Wuhan. *Nature* **584**, 420–424 (2020).
2. Y. Dong *et al*., Epidemiology of COVID-19 Among Children in China. *Pediatrics* **145**, e20200702 (2020).
3. C. Heneghan, J. Brassey, T. Jefferson, COVID-19: What proportion are asymptomatic? https://www.cebm.net/covid-19/covid-19-what-proportion-are-asymptomatic/. Accessed 29 May 2020.
4. M. Day, Covid-19: Four fifths of cases are asymptomatic, China figures indicate. *BMJ* **369**, m1375 (2020).
5. S. Mayor, Covid-19: Nine in 10 pregnant women with infection when admitted for delivery are asymptomatic, small study finds. *BMJ* **369**, m1485 (2020).
6. Centers for Disease Control and Prevention, COVID-19 Pandemic Planning Scenarios. https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios-archive/planning-scenarios-2020-05-20.pdf. Accessed 18 August 2020.
7. J. He, Y. Guo, R. Mao, J. Zhang, Proportion of asymptomatic coronavirus disease 2019 (COVID-19): A systematic review and meta-analysis. *J. Med. Virol.* **93**, 820–830 (2021).
8. O. Byambasuren *et al*., Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: Systematic review and meta-analysis. *J. Assoc. Med. Microbiol. Infect. Dis. Can.*, **5**, 223–234 (2020).
9. D. Buitrago-Garcia *et al*., Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: A living systematic review and meta-analysis. *PLoS Med.* **17**, e1003346 (2020).
10. D. F. Gudbjartsson *et al*., Spread of SARS-CoV-2 in the Icelandic population. *N. Engl. J. Med.* **382**, 2302–2315 (2020).
11. A. Kimball *et al*., Asymptomatic and presymptomatic SARS-CoV-2 infections in residents of a long-term care skilled nursing facility - King County, Washington, March 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 377–381 (2020).
12. M. M. Arons *et al*., Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. *N. Engl. J. Med.* **382**, 2081–2090 (2020).
13. S. Hoehl *et al*., Evidence of SARS-CoV-2 infection in returning travelers from Wuhan, China. *N. Engl. J. Med.* **382**, 1278–1280 (2020).
14. Q. M. L. Thi *et al*., Severe acute respiratory syndrome coronavirus 2 shedding by travelers, Vietnam, 2020. *Emerging. Infect. Dis. J.* **26**, 1624–1626 (2020).
15. R. Zhou *et al*., Viral dynamics in asymptomatic patients with COVID-19. *Int. J. Infect. Dis.* **96**, 288–290 (2020).
16. S. M. Moghadas *et al*., The implications of silent transmission for the control of COVID-19 outbreaks. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 17513–17515 (2020).
17. S. Almazeedi *et al*., Characteristics, risk factors and outcomes among the first consecutive 1096 patients diagnosed with COVID-19 in Kuwait. *EClinicalMedicine* **24**, 100448 (2020).
18. M. Al-Qahtani *et al*., The prevalence of asymptomatic and symptomatic COVID-19 in a cohort of quarantined subjects. *Int. J. Infect. Dis.* **102**, 285–288 (2021).
19. H. O. Al-Shamsi, E. A. Coomes, K. Aldhaheri, S. Alrawi, Serial screening for COVID-19 in asymptomatic patients receiving anticancer therapy in the United Arab Emirates. *JAMA Oncol.* **7**, 129–131 (2021).
20. D. Alsharrah *et al*., Clinical characteristics of pediatric SARS-CoV-2 infection and coronavirus disease 2019 (COVID-19) in Kuwait. *J. Med. Virol.* **93**, 3246–3250 (2021).
21. G. R. Alvarado *et al*., Symptom characterization and outcomes of sailors in isolation after a COVID-19 outbreak on a US aircraft carrier. *JAMA Netw. Open* **3**, e2020981 (2020).
22. P. An, P. Song, Y. Wang, B. Liu, Asymptomatic patients with novel coronavirus disease (COVID-19). *Balkan Med. J.* **37**, 229–230 (2020).
23. P. Anand *et al*., Clinical profile, viral load, management and outcome of neonates born to COVID 19 positive mothers: A tertiary care centre experience from India. *Eur. J. Pediatr.* **180**, 547–559 (2021).
24. Y. Arima *et al*., Severe acute respiratory syndrome coronavirus 2 infection among returnees to Japan from Wuhan, China, 2020. *Emerg. Infect. Dis.* **26**, 1596–1600 (2020).
25. M. S. Alshahrani *et al*., Prevalence of the SARS-CoV-2 infection among post-quarantine healthcare workers. *J. Multidiscip. Healthc.* **13**, 1927–1936 (2020).
26. A. Aslam *et al*., SARS CoV-2 surveillance and exposure in the perioperative setting with universal testing and personal protective equipment (PPE) policies. *Clin. Infect. Dis.*, 10.1093/cid/ciaa1607 (2020).
27. E. Atalla *et al*., Clinical presentation, course, and risk factors associated with mortality in a severe outbreak of COVID-19 in Rhode Island, USA, April-June 2020. *Pathogens* **10**, 8 (2020).
28. S. H. Bae *et al*., Asymptomatic transmission of SARS-CoV-2 on evacuation flight. *Emerg. Infect. Dis.* **26**, 2705–2708 (2020).
29. W. R. Bender, A. Hirshberg, P. Coutifaris, A. L. Acker, S. K. Srinivas, Universal testing for severe acute respiratory syndrome coronavirus 2 in 2 Philadelphia hospitals: Carrier prevalence and symptom development over 2 weeks. *Am. J. Obstet. Gynecol. MFM* **2**, 100226 (2020).

30. H. Blain *et al*., Efficacy of a test-retest strategy in residents and health care personnel of a nursing home facing a COVID-19 outbreak. *J. Am. Med. Dir. Assoc.* **21**, 933–936 (2020).
31. H. Blain *et al*., Atypical clinical presentation of COVID-19 infection in residents of a long-term care facility. *Eur. Geriatr. Med.* **11**, 1085–1088 (2020).
32. J. F.-W. Chan *et al*., A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: A study of a family cluster. *Lancet* **395**, 514–523 (2020).
33. L. Chang, L. Zhao, H. Gong, L. Wang, L. Wang, Severe acute respiratory syndrome coronavirus 2 RNA detected in blood donations. *Emerg. Infect. Dis.* **26**, 1631–1633 (2020).
34. N. Van Vinh Chau *et al*., The natural history and transmission potential of asymptomatic SARS-CoV-2 infection. *bioRxiv* [Preprint] (2020). https://doi.org/10.1101/2020.04.27.20082347.
35. M. Chehrassan *et al*., Management of spine trauma in COVID-19 pandemic: A preliminary report. *Arch. Bone Jt. Surg.* **8**, 270–276 (2020).
36. L. Coppeta *et al*., Contact screening for healthcare workers exposed to patients with COVID-19. *Int. J. Environ. Res. Public Health* **17**, 9082 (2020).
37. M. A. Corcorran *et al*., Prolonged persistence of PCR-detectable virus during an outbreak of SARS-CoV-2 in an inpatient geriatric psychiatry unit in King County, Washington. *Am. J. Infect. Control* **49**, 293–298 (2021).
38. S. Cosma *et al*., The "scar" of a pandemic: Cumulative incidence of COVID-19 during the first trimester of pregnancy. *J. Med. Virol.* **93**, 537–540 (2021).
39. K. Danis *et al*., Cluster of coronavirus disease 2019 (COVID-19) in the French Alps, February 2020. *Clin. Infect. Dis.* **71**, 825–832 (2020).
40. A. V. Dora *et al*., Universal and serial laboratory testing for SARS-CoV-2 at a long-term care skilled nursing facility for Veterans - Los Angeles, California, 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 651–655 (2020).
41. W. Du *et al*., Clinical characteristics of COVID-19 in children compared with adults in Shandong Province, China. *Infection* **48**, 445–452 (2020).
42. W. Du *et al*., Persistence of SARS-CoV-2 virus RNA in feces: A case series of children. *J. Infect. Public Health* **13**, 926–931 (2020).
43. E. Eythorsson *et al*., Clinical spectrum of coronavirus disease 2019 in Iceland: Population based cohort study. *BMJ* **371**, m4529 (2020).
44. F. M. Fusco *et al*., COVID-19 among healthcare workers in a specialist infectious diseases setting in Naples, Southern Italy: Results of a cross-sectional surveillance study. *J. Hosp. Infect.* **105**, 596–600 (2020).
45. Y. Gao *et al*., A cluster of the corona virus disease 2019 caused by incubation period transmission in Wuxi, China. *J. Infect.* **80**, 666–670 (2020).
46. A. Gonfiotti *et al*., Clinical courses and outcomes of five patients with primary lung cancer surgically treated while affected by severe acute respiratory syndrome coronavirus 2. *Eur. J. Cardiothorac. Surg.* **58**, 598–604 (2020).
47. C. G. Grijalva *et al*., Transmission of SARS-COV-2 infections in households - Tennessee and Wisconsin, April-September 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 1631–1634 (2020).
48. N. Gupta *et al*., Transmission of SARS-CoV-2 infection by children: A study of contacts of index paediatric cases in India. *J. Trop. Pediatr.* **67**, fmaa081 (2021).
49. M. S. Han *et al*., Clinical characteristics and viral RNA detection in children with coronavirus disease 2019 in the Republic of Korea. *JAMA Pediatr.* **175**, 73–80 (2021).
50. S. Harada, *et al*., Control of a nosocomial outbreak of COVID-19 in a university hospital. *Open Forum Infect. Dis.* **7**, ofaa512 (2020).
51. I. Hasanoglu *et al*., Higher viral loads in asymptomatic COVID-19 patients might be the invisible part of the iceberg. *Infection* **49**, 117–126 (2021).
52. N. Hcini *et al*., Maternal, fetal and neonatal outcomes of large series of SARS-CoV-2 positive pregnancies in peripartum period: A single-center prospective comparative study. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **257**, 11–18 (2021).
53. X. He *et al*., Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* **26**, 672–675 (2020).
54. C. A. Hogan *et al*., Large-scale testing of asymptomatic healthcare personnel for severe acute respiratory syndrome coronavirus 2. *Emerg. Infect. Dis.* **27**, 250–254 (2021).
55. Z. Hu *et al*., Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China. *Sci. China Life Sci.* **63**, 706–711 (2020).
56. C.-Z. Hua *et al*., Epidemiological features and viral shedding in children with SARS-CoV-2 infection. *J. Med. Virol.* **92**, 2804–2812 (2020).
57. L. Huang *et al*., Initial CT imaging characters of an imported family cluster of COVID-19. *Clin. Imaging* **65**, 78–81 (2020).
58. K. Huang *et al*., A retrospective analysis of the epidemiology, clinical manifestations, and imaging characteristics of familial cluster-onset COVID-19. *Ann. Transl. Med.* **8**, 747 (2020).
59. Q. Huang *et al*., Asymptomatic COVID-19 infection in patients with cancer at a cancer-specialized hospital in Wuhan, China - Preliminary results. *Eur. Rev. Med. Pharmacol. Sci.* **24**, 9760–9764 (2020).
60. I. F.-N. Hung *et al*., SARS-CoV-2 shedding and seroconversion among passengers quarantined after disembarking a cruise ship: A case series. *Lancet Infect. Dis.* **20**, 1051–1060 (2020).
61. O. R. Ibrahim *et al*., COVID-19 in children: A case series from Nigeria. *Pan Afr. Med. J.* **35**, 53 (2020).
62. T. H. Jeong *et al*., Real asymptomatic SARS-CoV-2 infection might be rare: Importance of careful interviews and follow-up. *J. Korean Med. Sci.* **35**, e333 (2020).
63. X.-L. Jiang *et al*., Transmission potential of asymptomatic and paucisymptomatic severe acute respiratory syndrome coronavirus 2 infections: A 3-family cluster study in China. *J. Infect. Dis.* **221**, 1948–1952 (2020).
64. Y. Jiang *et al*., Characteristics of a family cluster of Severe Acute Respiratory Syndrome Coronavirus 2 in Henan, China. *J. Infect.* **81**, e46–e48 (2020).
65. H. Jiang *et al*., Clinical features, laboratory findings and persistence of virus in 10 children with coronavirus disease 2019 (COVID-19). *Biomed. J.* **44**, 94–100 (2021).
66. N. K. Jones *et al*., Effective control of SARS-CoV-2 transmission between healthcare workers during a period of diminished community prevalence of COVID-19. *eLife* **9**, e59391 (2020).
67. R. K. Joshi, R. K. Ray, S. Adhya, V. P. S. Chauhan, S. Pani, Spread of COVID-19 by asymptomatic cases: Evidence from military quarantine facilities. *BMJ Mil. Health* **167**, 217–218 (2021).
68. J. Jung *et al*., Investigation of a nosocomial outbreak of coronavirus disease 2019 in a paediatric ward in South Korea: Successful control by early detection and extensive contact tracing with testing. *Clin. Microbiol. Infect.* **26**, 1574–1575 (2020).
69. Y. Kang *et al*., A retrospective view of pediatric cases infected with SARS-CoV-2 of a middle-sized city in mainland China. *Medicine (Baltimore)* **99**, e23797 (2020).
70. L. Karout *et al*., COVID-19 prevalence, risk perceptions, and preventive behavior in asymptomatic Latino population: A cross-sectional study. *Cureus* **12**, e10707 (2020).
71. S. P. Kennelly *et al*., Asymptomatic carriage rates and case fatality of SARS-CoV-2 infection in residents and staff in Irish nursing homes. *Age Ageing* **50**, 49–54 (2021).
72. S. C. Kirshblum *et al*., Screening testing for SARS-CoV-2 upon admission to rehabilitation hospitals in a high COVID-19 prevalence community. *PM R.* **12**, 1009–1014 (2020).
73. B. R. Kittang *et al*., Outbreak of COVID-19 at three nursing homes in Bergen. *Tidsskr. Nor. Laegeforen.*, 10.4045/tidsskr.20.0405 (2020).
74. S. Kutsuna, *et al*., SARS-CoV-2 screening test for Japanese returnees from Wuhan, China, January 2020. *Open Forum Infect. Dis.* **7**, ofaa243(2020).
75. S. N. Ladhani *et al*., Investigation of SARS-CoV-2 outbreaks in six care homes in London, April 2020. *EClinicalMedicine* **26**, 100533 (2020).
76. J. Y. Lee *et al*., Epidemiological and clinical characteristics of coronavirus disease 2019 in Daegu, South Korea. *Int. J. Infect. Dis.* **98**, 462–466 (2020).
77. W. Li *et al*., Virus shedding dynamics in asymptomatic and mildly symptomatic patients infected with SARS-CoV-2. *Clin. Microbiol. Infect.* **26**, 1556.e1–1556.e6 (2020).
78. J. Li *et al*., Comparative analysis of symptomatic and asymptomatic SARS-CoV-2 infection in children. *Ann. Acad. Med. Singap.* **49**, 530–537 (2020).
79. J. Liao *et al*., Epidemiological and clinical characteristics of COVID-19 in adolescents and young adults. *Innovation (N Y)* **1**, 100001 (2020).
80. Z. Liu *et al*., Investigation of a family cluster outbreak of COVID-19 indicates the necessity of CT screening for asymptomatic family members in close contact with confirmed patients. *J. Thorac. Dis.* **12**, 3673–3681 (2020).
81. A. Lombardi *et al*., Characteristics of 1573 healthcare workers who underwent nasopharyngeal swab testing for SARS-CoV-2 in Milan, Lombardy, Italy. *Clin. Microbiol. Infect.* **26**, 1413.e9–1413.e13 (2020).
82. L. Luo *et al*., Contact settings and risk for transmission in 3410 close contacts of patients with COVID-19 in Guangzhou, China: A prospective cohort study. *Ann. Intern. Med.* **173**, 879–887 (2020).
83. Y. Luo *et al*., Asymptomatic SARS-CoV-2 infection in household contacts of a healthcare provider, Wuhan, China. *Emerg. Infect. Dis.* **26**, 1930–1933 (2020).
84. T. Lytras *et al*., High prevalence of SARS-CoV-2 infection in repatriation flights to Greece from three European countries. *J. Travel Med.* **27**, taaa054 (2020).
85. Y. Ma *et al*., Characteristics of asymptomatic patients with SARS-CoV-2 infection in Jinan, China. *Microbes Infect.* **22**, 212–217 (2020).
86. K. Macartney *et al*., Transmission of SARS-CoV-2 in Australian educational settings: A prospective cohort study. *Lancet Child Adolesc. Health* **4**, 807–816 (2020).
87. M. L. Martinez-Fierro *et al*., The role of close contacts of COVID-19 patients in the SARS-CoV-2 transmission: An emphasis on the percentage of nonevaluated positivity in Mexico. *Am. J. Infect. Control* **49**, 15–20 (2021).
88. M. A. Merza, A. A. Haleem Al Mezori, H. M. Mohammed, D. M. Abdulah, COVID-19 outbreak in Iraqi Kurdistan: The first report characterizing epidemiological, clinical, laboratory, and radiological findings of the disease. *Diabetes Metab. Syndr.* **14**, 547–554 (2020).
89. K. J. Meyers *et al*., A cross-sectional community-based observational study of asymptomatic SARS-CoV-2 prevalence in the greater Indianapolis area. *J. Med. Virol.* **92**, 2874–2879 (2020).
90. R. Migisha *et al*., Early cases of SARS-CoV-2 infection in Uganda: Epidemiology and lessons learned from risk-based testing approaches - March-April 2020. *Global. Health* **16**, 114 (2020).
91. O.-T. Ng *et al*., SARS-CoV-2 infection among travelers returning from Wuhan, China. *N. Engl. J. Med.* **382**, 1476–1478 (2020).
92. H. Njuguna *et al*., Serial laboratory testing for SARS-CoV-2 infection among incarcerated and detained persons in a correctional and detention facility - Louisiana, April-May 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 836–840 (2020).
93. S. Y. Park *et al*., Coronavirus disease outbreak in call center, South Korea. *Emerg. Infect. Dis.* **26**, 1666–1670 (2020).
94. J. H. Park, J. H. Jang, K. Lee, S. J. Yoo, H. Shin, COVID-19 outbreak and presymptomatic transmission in pilgrim travelers who returned to Korea from Israel. *J. Korean Med. Sci.* **35**, e424 (2020).
95. V. C. Passarelli *et al*., Asymptomatic SARS-CoV-2 infections in hospitalized patients. *Infect. Control Hosp. Epidemiol.*, 10.1017/ice.2020.441 (2020).
96. V. C. Passarelli *et al*., Asymptomatic COVID-19 in hospital visitors: The underestimated potential of viral shedding. *Int. J. Infect. Dis.* **102**, 412–414 (2021).
97. M. C. Patel *et al*., Asymptomatic SARS-CoV-2 infection and COVID-19 mortality during an outbreak investigation in a skilled nursing facility. *Clin. Infect. Dis.* **71**, 2920–2926 (2020).
98. A. Pavli *et al*., A cluster of COVID-19 in pilgrims to Israel. *J. Travel Med.* **27**, taaa102 (2020).

99. P. Q. Thai, *et al*., The first 100 days of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) control in Vietnam. *Clin. Infect. Dis.* **72**, e334–e342 (2021).
100. M. M. Plucinski *et al*., COVID-19 in Americans aboard the Diamond Princess cruise ship. *Clin. Infect. Dis.* **72**, e448–e457 (2021).
101. W. A. Pongpirul *et al*., Clinical characteristics of patients hospitalized with coronavirus disease, Thailand. *Emerg. Infect. Dis.* **26**, 1580–1585 (2020).
102. C. A. J. Puylaert, *et al*., Yield of screening for COVID-19 in asymptomatic patients prior to elective or emergency surgery using chest CT and RT-PCR (SCOUT): Multicenter study. *Ann. Surg.* **272**, 919–924 (2020).
103. G. Qian *et al*., COVID-19 transmission within a family cluster by presymptomatic carriers in China. *Clin. Infect. Dis.* **71**, 861–862 (2020).
104. W. Qin *et al*., The descriptive epidemiology of coronavirus disease 2019 during the epidemic period in Lu'an, China: Achieving limited community transmission using proactive response strategies. *Epidemiol. Infect.* **148**, e132 (2020).
105. C. Qiu *et al*., Transmission and clinical characteristics of coronavirus disease 2019 in 104 outside-Wuhan patients, China. *J. Med. Virol.* **92**, 2027–2035 (2020).
106. M. Ralli, A. Morrone, A. Arcangeli, L. Ercoli, Asymptomatic patients as a source of transmission of COVID-19 in homeless shelters. *Int. J. Infect. Dis.* **103**, 243–245 (2021).
107. M. L. Reale *et al*., SARS-CoV-2 infection in cancer patients: A picture of an Italian *Onco-Covid* unit. *Front. Oncol.* **10**, 1722 (2020).
108. A. Rincón *et al*., The keys to control a COVID-19 outbreak in a haemodialysis unit. *Clin. Kidney J.* **13**, 542–549 (2020).
109. F. Rivera *et al*., Prevalence of SARS-CoV-2 asymptomatic infections in two large academic health systems in Wisconsin. *Clin. Infect. Dis.*, 10.1093/cid/ciaa1225 (2020).
110. L. Rivett *et al*., Screening of healthcare workers for SARS-CoV-2 highlights the role of asymptomatic carriage in COVID-19 transmission. *eLife* **9**, e58728 (2020).
111. K. Saeed *et al*., Investigations, actions and learning from an outbreak of SARS-CoV-2 infection among healthcare workers in the United Kingdom. *J. Infect. Prev.* **22**, 156–161 (2021).
112. R. R. Ossami Saidy, B. Globke, J. Pratschke, W. Schoening, D. Eurich, Successful implementation of preventive measures leads to low relevance of SARS-CoV-2 in liver transplant patients: Observations from a German outpatient department. *Transpl. Infect. Dis.* **22**, e13363 (2020).
113. V. Schwierzeck *et al*., First reported nosocomial outbreak of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in a pediatric dialysis unit. *Clin. Infect. Dis.* **72**, 265–270 (2021).
114. K. C. See *et al*., COVID-19: Four paediatric cases in Malaysia. *Int. J. Infect. Dis.* **94**, 125–127 (2020).
115. A. P. Senanayake *et al*., Features of Covid-19 patients detected during community screening: A study from a rural hospital in Sri Lanka. *Ceylon Med. J.* **65**, 67 (2020).
116. V. Setiawaty *et al*., The identification of first COVID-19 cluster in Indonesia. *Am. J. Trop. Med. Hyg.* **103**, 2339–2342 (2020).
117. J. Shen *et al*., Characteristics of nosocomial infections in children screened for SARS-CoV-2 infection in China. *Med. Sci. Monit.* **26**, e928835 (2020).
118. S. M. Shi *et al*., Risk factors, presentation, and course of coronavirus disease 2019 in a large, academic long-term care facility. *J. Am. Med. Dir. Assoc.* **21**, 1378–1383.e1 (2020).
119. H. Son *et al*., Epidemiological characteristics of and containment measures for COVID-19 in Busan, Korea. *Epidemiol. Health* **42**, e2020035 (2020).
120. W. Song *et al*., Clinical features of pediatric patients with coronavirus disease (COVID-19). *J. Clin. Virol.* **127**, 104377 (2020).
121. A. Soysal *et al*., Comparison of clinical and laboratory features and treatment options of 237 Comparison of clinical and laboratory features and treatment options of 237 symptomatic and asymptomatic children infected with SARS-CoV-2 in the early phase of the COVID-19 pandemic in Turkey. *Jpn. J. Infect. Dis.* **74**, 273–279 (2021).
122. A. D. Stock *et al*., COVID-19 infection among healthcare workers: Serological findings supporting routine testing. *Front. Med. (Lausanne)* **7**, 471 (2020).
123. H. J. Suh *et al*., Clinical characteristics of COVID-19: Clinical dynamics of mild severe acute respiratory syndrome coronavirus 2 infection detected by early active surveillance. *J. Korean Med. Sci.* **35**, e297 (2020).
124. S. Szegedi, W. Huf, K. Miháltz, P. V. Vécsei-Marlovits, Prevalence of SARS-CoV-2 infection in patients presenting for intravitreal injection. *Spektrum Der Augenheilkunde* **35** (70), 74 (2020).
125. S. Tabata *et al*., Clinical characteristics of COVID-19 in 104 people with SARS-CoV-2 infection on the Diamond Princess cruise ship: A retrospective analysis. *Lancet Infect. Dis.* **20**, 1043–1050 (2020).
126. Y.-P. Tan *et al*., Epidemiologic and clinical characteristics of 10 children with coronavirus disease 2019 in Changsha, China. *J. Clin. Virol.* **127**, 104353 (2020).
127. J. Taylor *et al*., Serial testing for SARS-CoV-2 and virus whole genome sequencing inform infection risk at two skilled nursing facilities with COVID-19 outbreaks - Minnesota, April-June 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 1288–1295 (2020).
128. H. Temel *et al*., Evaluation of the clinical features of 81 patients with COVID-19: An unpredictable disease in children. *J. Pediatr. Infect. Dis.* **16**, 47–52 (2021).
129. J. W. Vivian Thangaraj *et al*., A cluster of SARS-CoV-2 infection among Italian tourists visiting India, March 2020. *Indian J. Med. Res.* **151**, 438–443 (2020).
130. S. L. Thiel *et al*., Flattening the curve in 52 days: Characterisation of the COVID-19 pandemic in the Principality of Liechtenstein - An observational study. *Swiss Med. Wkly.* **150**, w20361 (2020).
131. Z.-D. Tong *et al*., Potential presymptomatic transmission of SARS-CoV-2, Zhejiang Province, China, 2020. *Emerg. Infect. Dis.* **26**, 1052–1054 (2020).
132. T.-P. Tsou *et al*., Epidemiology of the first 100 cases of COVID-19 in Taiwan and its implications on outbreak control. *J. Formos. Med. Assoc.* **119**, 1601–1607 (2020).
133. A. Wadhwa *et al*., Identification of presymptomatic and asymptomatic cases using cohort-based testing approaches at a large correctional facility - Chicago, Illinois, USA, May 2020. *Clin. Infect. Dis.* **72**, e128–e135 (2021).
134. Y. Wang *et al*., Characterization of an asymptomatic cohort of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) infected individuals outside of Wuhan, China. *Clin. Infect. Dis.* **71**, 2132–2138 (2020).
135. G. Wang *et al*., Infection, screening, and psychological stress of health care workers with COVID-19 in a non-frontline clinical department. *Disaster Med. Public Health Prep.*, 10.1017/dmp.2020.428 (2020).
136. Y. M. Wi *et al*., Response system for and epidemiological features of COVID-19 in Gyeongsangnam-do Province in South Korea. *Clin. Infect. Dis.* **72**, 661–667 (2021).
137. J. Wong *et al*., High proportion of asymptomatic and presymptomatic COVID-19 infections in air passengers to Brunei. *J. Travel Med.* **27**, taaa066 (2020).
138. J. Wong, S. A. Jamaludin, M. F. Alikhan, L. Chaw, Asymptomatic transmission of SARS-CoV-2 and implications for mass gatherings. *Influenza Other Respir. Viruses* **14**, 596–598 (2020).
139. J. Wu *et al*., Household transmission of SARS-CoV-2, Zhuhai, China, 2020. *Clin. Infect. Dis.* **71**, 2099–2108 (2020).
140. S. Wu *et al*., Understanding factors influencing the length of hospital stay among non-severe COVID-19 patients: A retrospective cohort study in a Fangcang shelter hospital. *PLoS One* **15**, e0240959 (2020).
141. A. Xi *et al*., Epidemiological and clinical characteristics of discharged patients infected with SARS-CoV-2 on the Qinghai Plateau. *J. Med. Virol.* **92**, 2528–2535 (2020).
142. F. Xiong *et al*., Clinical characteristics of and medical interventions for COVID-19 in hemodialysis patients in Wuhan, China. *J. Am. Soc. Nephrol.* **31**, 1387–1397 (2020).
143. M.-C. Yang *et al*., A three-generation family cluster with COVID-19 infection: Should quarantine be prolonged? *Public Health* **185**, 31–33 (2020).
144. R. Yang, X. Gui, Y. Xiong, Comparison of clinical characteristics of patients with asymptomatic vs symptomatic coronavirus disease 2019 in Wuhan, China. *JAMA Netw. Open* **3**, e2010182 (2020).
145. N. Yang *et al*., In-flight transmission cluster of COVID-19: A retrospective case series. *Infect. Dis. (Lond.)* **52**, 891–901 (2020).
146. K. Yau *et al*., COVID-19 outbreak in an urban hemodialysis unit. *Am. J. Kidney Dis.* **76**, 690–695.e1 (2020).
147. F. Ye *et al*., Delivery of infection from asymptomatic carriers of COVID-19 in a familial cluster. *Int. J. Infect. Dis.* **94**, 133–138 (2020).
148. J. Zhang, S. Tian, J. Lou, Y. Chen, Familial cluster of COVID-19 infection from an asymptomatic. *Crit. Care* **24**, 119 (2020).
149. W. Zhang *et al*., Secondary transmission of coronavirus disease from presymptomatic persons, China. *Emerg. Infect. Dis.* **26**, 1924–1926 (2020).
150. H.-J. Zhang *et al*., Asymptomatic and symptomatic SARS-CoV-2 infections in close contacts of COVID-19 patients: A seroepidemiological study. *Clin. Infect. Dis.*, 10.1093/cid/ciaa771 (2020).
151. H. Zhang, R. Chen, J. Chen, B. Chen, COVID-19 transmission within a family cluster in Yancheng, China. *Front. Med. (Lausanne)* **7**, 387 (2020).
152. H. Zhang *et al*., A multi-family cluster of COVID-19 associated with asymptomatic and pre-symptomatic transmission in Jixi City, Heilongjiang, China, 2020. *Emerg. Microbes Infect.* **9**, 2509–2514 (2020).
153. D. Zhao *et al*., Asymptomatic infection by SARS-CoV-2 in healthcare workers: A study in a large teaching hospital in Wuhan, China. *Int. J. Infect. Dis.* **99**, 219–225 (2020).
154. X. Zheng *et al*., Asymptomatic patients and asymptomatic phases of Coronavirus Disease 2019 (COVID-19): A population-based surveillance study. *Natl. Sci. Rev.* **7**, 1527–1539 (2020).
155. C. Olmos *et al*., SARS-CoV-2 infection in asymptomatic healthcare workers at a clinic in Chile. *PLoS One* **16**, e0245913 (2021).
156. S. C. Roberts *et al*., Mass severe acute respiratory coronavirus 2 (SARS-CoV-2) testing of asymptomatic healthcare personnel - ERRATUM. *Infect. Control Hosp. Epidemiol.* **42**, 625–626 (2021).
157. J. Tan-Loh, B. M. K. Cheong, A descriptive analysis of clinical characteristics of COVID-19 among healthcare workers in a district specialist hospital. *Med. J. Malaysia* **76**, 24–28 (2021).
158. M. Goldenfeld *et al*., Characteristics of clinically asymptomatic patients with SARS-CoV-2 infections, case series. *Prehosp. Disaster Med.* **36**, 125–128 (2021).
159. R. Green *et al*., COVID-19 testing in outbreak-free care homes: What are the public health benefits? *J. Hosp. Infect.* **111**, 89–95 (2021).
160. M. Krone, A. Noffz, E. Richter, U. Vogel, M. Schwab, Control of a COVID-19 outbreak in a nursing home by general screening and cohort isolation in Germany, March to May 2020. *Euro Surveill.* **26**, 2001365 (2021).
161. S. Krüger *et al*., Performance and feasibility of universal PCR admission screening for SARS-CoV-2 in a German tertiary care hospital. *J. Med. Virol.* **93**, 2890–2898 (2021).
162. R. J. Reid, L. Rosella, N. Milijasevic, L. N. Small, Mass testing for asymptomatic COVID-19 infection among health care workers at a large Canadian hospital. *J. Assoc. Med. Microbiol. Infect. Dis. Can.* **5**, 245–250 (2020).
163. A. Choudhury *et al*., COVID-19 in liver transplant recipients—A series with successful recovery. *J. Clin. Transl. Hepatol.* **8**, 467–473 (2020).
164. K. J. Meyers *et al*., Follow-up of SARS-CoV-2 positive subgroup from the asymptomatic novel CORonavirus iNFection study. *J. Med. Virol.* **93**, 2925–2931 (2021).
165. M. Cruz-Lemini *et al*. Obstetric outcomes of SARS-CoV-2 infection in asymptomatic pregnant women. *Viruses* **13**, 112 (2021).
166. H. E. Abraha *et al*., Clinical features and risk factors associated with morbidity and mortality among patients with COVID-19 in northern Ethiopia. *Int. J. Infect. Dis.* **105**, 776–783 (2021).
167. K. J. Beiting *et al*., Management and outcomes of a COVID-19 outbreak in a nursing home with predominantly Black residents. *J. Am. Geriatr. Soc.* **69**, 1155–1165 (2021).
168. H. Blain *et al*., Atypical symptoms, SARS-CoV-2 test results and immunisation rates in 456 residents from eight nursing homes facing a COVID-19 outbreak. *Age Ageing* **50**, 641–648 (2021).

169. J. Chan *et al*., COVID-19 in the New York City jail system: Epidemiology and health care response, March-April 2020. *Public Health Rep.* **136**, 375–383 (2021).
170. N. Gupta *et al*., Clinical profile and outcomes of asymptomatic vs. symptomatic travellers diagnosed with COVID-19: An observational study from a coastal town in South India. *Drug Discov. Ther.* **15**, 1–8 (2021).
171. M. J. Gutman *et al*., What was the prevalence of COVID-19 in asymptomatic patients undergoing orthopaedic surgery in one large United States City Mid-pandemic? *Clin. Orthop. Relat. Res.* **479**, 1691–1699 (2021).
172. N. Handal *et al*., Comparison of SARS-CoV-2 infections in healthcare workers with high and low exposures to Covid-19 patients in a Norwegian University Hospital. *Infect. Dis. (Lond.)* **53**, 420–429 (2021).
173. S. Lamichhane, S. Gupta, G. Akinjobi, N. Ndubuka, Familial cluster of asymptomatic COVID-19 cases in a First Nation community in Northern Saskatchewan, Canada. *Can. Commun. Dis. Rep.* **47**, 94–96 (2021).
174. P. Liu *et al*., Epidemiological and clinical features in patients with coronavirus disease 2019 outside of Wuhan, China: Special focus in asymptomatic patients. *PLoS Negl. Trop. Dis.* **15**, e0009248 (2021).
175. J. N. Malagón-Rojas, M. Mercado, C. P. Gómez-Rendón, SARS-CoV-2 and work-related transmission: Results of a prospective cohort of airport workers, 2020. *Rev. Bras. Med. Trab.* **18**, 371–380 (2021).
176. M. Marks *et al*., Transmission of COVID-19 in 282 clusters in Catalonia, Spain: A cohort study. *Lancet Infect. Dis.* **21**, 629–636 (2021).
177. D. Nicolás *et al*., A prospective cohort of SARS-CoV-2-infected health care workers: Clinical characteristics, outcomes, and follow-up strategy. *Open Forum Infect. Dis.* **8**, ofaa592 (2021).
178. U. P. Patil, P. Krishnan, S. Abudinen-Vasquez, S. Maru, L. Noble, Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) positive newborns of COVID-19 mothers after dyad-care: A case series. *Cureus* **13**, e12528 (2021).
179. S. Rajme-López *et al*., Large-scale screening for severe acute respiratory coronavirus virus 2 (SARS-CoV-2) among healthcare workers: Prevalence and risk factors for asymptomatic and pauci-symptomatic carriers, with emphasis on the use of personal protective equipment (PPE). *Infect. Control Hosp. Epidemiol.*, 10.1017/ice.2021.68 (2021).
180. R. Ren *et al*., Asymptomatic SARS-CoV-2 infections among persons entering China from April 16 to October 12, 2020. *JAMA* **325**, 489–492 (2021).
181. R. Sharma *et al*., Perinatal outcome and possible vertical transmission of coronavirus disease 2019: Experience from North India. *Clin. Exp. Pediatr.* **64**, 239–246 (2021).
182. B. Stessel *et al*., Evaluation of a comprehensive pre-procedural screening protocol for COVID-19 in times of a high SARS CoV-2 prevalence: A prospective cross-sectional study. *Ann. Med.* **53**, 337–344 (2021).
183. J. K. Tan *et al*., The prevalence and clinical significance of Presymptomatic COVID-19 patients: How we can be one step ahead in mitigating a deadly pandemic. *BMC Infect. Dis.* **21**, 249 (2021).
184. S. Abeysuriya *et al*., Universal screening for SARS-CoV-2 in pregnant women at term admitted to an East London maternity unit. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **252**, 444–446 (2020).
185. F. Adorni *et al*., Self-reported symptoms of SARS-CoV-2 infection in a non-hospitalized population: Results from the large Italian web-based EPICOVID19 cross-sectional survey. *JMIR Public Health Surveill.* **6**, e21866 (2020).
186. S. Aherfi, P. Gautret, H. Chaudet, D. Raoult, B. La Scola, Clusters of COVID-19 associated with Purim celebration in the Jewish community in Marseille, France, March 2020. *Int. J. Infect. Dis.* **100**, 88–94 (2020).
187. M. Albalate *et al*., High prevalence of asymptomatic COVID-19 in hemodialysis. Daily learning during first month of COVID-19 pandemic. *Nefrologia (Engl. Ed.)* **40**, 279–286 (2020).
188. M. Andrikopoulou *et al*., Symptoms and critical illness among obstetric patients with Coronavirus Disease 2019 (COVID-19) infection. *Obstet. Gynecol.* **136**, 291–299 (2020).
189. T. P. Baggett, H. Keyes, N. Sporn, J. M. Gaeta, Prevalence of SARS-CoV-2 infection in residents of a large homeless shelter in Boston. *JAMA* **323**, 2191–2192 (2020).
190. Y. Bai *et al*., Presumed asymptomatic carrier transmission of COVID-19. *JAMA* **323**, 1406–1407 (2020).
191. S. L. Bai *et al*., Analysis of the first cluster of cases in a family of novel coronavirus pneumonia in Gansu Province. *Zhonghua Yu Fang Yi Xue Za Zhi* **54**, E005 (2020).
192. A. S. Berghoff *et al*., SARS-CoV-2 testing in patients with cancer treated at a tertiary care hospital during the COVID-19 pandemic. *J. Clin. Oncol.* **38**, 3547–3554 (2020).
193. Q. Bi *et al*., Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: A retrospective cohort study. *Lancet Infect. Dis.* **20**, 911–919 (2020).
194. M. J. Blitz *et al*., Universal testing for coronavirus disease 2019 in pregnant women admitted for delivery: Prevalence of peripartum infection and rate of asymptomatic carriers at four New York hospitals within an integrated healthcare system. *Am. J. Obstet. Gynecol. MFM* **2**, 100169 (2020).
195. T. J. Blumberg *et al*., Universal screening for COVID-19 in children undergoing orthopaedic surgery: A multicenter report. *J. Pediatr. Orthop.* **40**, e990–e993 (2020).
196. S. Brandstetter *et al*., Symptoms and immunoglobulin development in hospital staff exposed to a SARS-CoV-2 outbreak. *Pediatr. Allergy Immunol.* **31**, 841–847 (2020).
197. K. H. Campbell *et al*., Prevalence of SARS-CoV-2 among patients admitted for childbirth in Southern Connecticut. *JAMA* **323**, 2520–2522 (2020).
198. R. Cardona-Hernandez *et al*., Children and youth with diabetes are not at increased risk for hospitalization due to COVID-19. *Pediatr. Diabetes* **22**, 202–206 (2021).
199. C. Carroll *et al*., Routine testing of close contacts of confirmed COVID-19 cases—National COVID-19 contact management programme, Ireland, May to August 2020. *Public Health* **190**, 147–151 (2021).
200. A. M. Cattelan *et al*., An integrated strategy for the prevention of SARS-CoV-2 infection in healthcare workers: A prospective observational study. *Int. J. Environ. Res. Public Health* **17**, 5785 (2020).
201. M. C. Chang, W.-S. Seo, D. Park, J. Hur, Analysis of SARS-CoV-2 screening clinic (including drive-through system) data at a single university Hospital in South Korea from 27 January 2020 to 31 March 2020 during the COVID-19 outbreak. *Healthcare (Basel)* **8**, 145 (2020).
202. N. Chekhlabi *et al*., The epidemiological and clinical profile of COVID-19 in children: Moroccan experience of the Cheikh Khalifa University Center. *Pan Afr. Med. J.* **35**, 57 (2020).
203. P. Chen *et al*., Clinical and demographic characteristics of cluster cases and sporadic cases of Coronavirus Disease 2019 (COVID-19) in 141 patients in the main district of Chongqing, China, between January and February 2020. *Med. Sci. Monit.* **26**, e923985 (2020).
204. P. Chen *et al*., Epidemiological and clinical characteristics of 136 cases of COVID-19 in main district of Chongqing. *J. Formos. Med. Assoc.* **119**, 1180–1184 (2020).
205. Y. Chen *et al*., Epidemiological characteristics of infection in COVID-19 close contacts in Ningbo city. *Zhonghua Liu Xing Bing Xue Za Zhi* **41**, 667–671 (2020).
206. J. Chen *et al*., Potential transmission of SARS-CoV-2 on a flight from Singapore to Hangzhou, China: An epidemiological investigation. *Travel Med. Infect. Dis.* **36**, 101816 (2020).
207. G. Chen *et al*., Epidemiological analysis of 18 patients with COVID-19. *Eur. Rev. Med. Pharmacol. Sci.* **24**, 12522–12526 (2020).
208. H.-Y. Cheng *et al*., Contact tracing assessment of COVID-19 transmission dynamics in Taiwan and risk at different exposure periods before and after symptom onset. *JAMA Intern. Med.* **180**, 1156–1163 (2020).
209. C.-H. Chiu *et al*., Familial cluster of pneumonia and asymptomatic cases of COVID-19 in Taiwan. *J. Formos. Med. Assoc.* **119**, 1560–1561 (2020).
210. C. Clarke *et al*., High prevalence of asymptomatic COVID-19 infection in hemodialysis patients detected using serologic screening. *J. Am. Soc. Nephrol.* **31**, 1969–1975 (2020).
211. G. A. G. Crameri *et al*., Reduced maximal aerobic capacity after COVID-19 in young adult recruits, Switzerland, May 2020. *Euro Surveill.* **25**, 2001542 (2020).
212. F. D'Ambrosi *et al*., Management of gestational diabetes in women with a concurrent severe acute respiratory syndrome coronavirus 2 infection, experience of a single center in Northern Italy. *Int. J. Gynaecol. Obstet.* **152**, 335–338 (2020).
213. T. Stock da Cunha *et al*., The spectrum of clinical and serological features of COVID-19 in urban hemodialysis patients. *J. Clin. Med.* **9**, 2264 (2020).
214. J. H. D. Silva, E. C. D. Oliveira, T. Y. Hattori, E. R. S. D. Lemos, A. C. P. Terças-Trettel, Description of a COVID-19 cluster: isolation and testing in asymptomatic people as prevention strategies for local spread in Mato Grosso. *Epidemiol. Serv. Saude* **29**, e2020264 (2020).
215. M. C. S. Jesus *et al*., Family COVID-19 cluster analysis of an infant without respiratory symptoms. *Rev. Soc. Bras. Med. Trop.* **53**, e20200494 (2020).
216. A. Dhuyvetter, H. E. Cejtin, M. Adam, A. Patel, Coronavirus disease 2019 in pregnancy: The experience at an urban safety net hospital. *J. Community Health* **46**, 267–269 (2021).
217. P. Díaz-Corvillón *et al*., Routine screening for SARS CoV-2 in unselected pregnant women at delivery. *PLoS One* **15**, e0239887 (2020).
218. M. Donahue *et al*., Notes from the field: Characteristics of meat processing facility workers with confirmed SARS-CoV-2 infection - Nebraska, April-May 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 1020–1022 (2020).
219. H. Du *et al*., Clinical characteristics of 182 pediatric COVID-19 patients with different severities and allergic status. *Allergy* **76**, 510–532 (2021).
220. D. Wang *et al*., Clinical analysis of 31 cases of 2019 novel coronavirus infection in children from six provinces (autonomous region) of northern China. *Zhonghua Er Ke Za Zhi* **58**, 269–274 (2020).
221. A. G. Edlow *et al*., Assessment of maternal and neonatal SARS-CoV-2 viral load, transplacental antibody transfer, and placental pathology in pregnancies during the COVID-19 pandemic. *JAMA Netw. Open* **3**, e2030455 (2020).
222. C. Felice, G. L. Di Tanna, G. Zanus, U. Grossi, Impact of COVID-19 outbreak on healthcare workers in Italy: Results from a national E-survey. *J. Community Health* **45**, 675–683 (2020).
223. A. Gaur *et al*., Clinico-radiological presentation of COVID-19 patients at a tertiary care center at Bhilwara Rajasthan, India. *J. Assoc. Physicians India* **68**, 29–33 (2020).
224. I. T. Goldfarb *et al*., Universal SARS-CoV-2 testing on admission to the labor and delivery unit: Low prevalence among asymptomatic obstetric patients. *Infect. Control Hosp. Epidemiol.* **41**, 1095–1096 (2020).
225. X. Gong *et al*., Three infection clusters related with potential pre-symptomatic transmission of coronavirus disease (COVID-19), Shanghai, China, January to February 2020. *Euro Surveill.* **25**, 2000228 (2020).
226. N. S. N. Graham *et al*., SARS-CoV-2 infection, clinical features and outcome of COVID-19 in United Kingdom nursing homes. *J. Infect.* **81**, 411–419 (2020).
227. O. Grechukhina *et al*., Coronavirus disease 2019 pregnancy outcomes in a racially and ethnically diverse population. *Am. J. Obstet. Gynecol. MFM* **2**, 100246 (2020).
228. J. Greiner *et al*., Characteristics and mechanisms to control a COVID-19 outbreak on a leukemia and stem cell transplantation unit. *Cancer Med.* **10**, 237–246 (2021).
229. J. A. Gruskay *et al*., Universal testing for COVID-19 in essential orthopaedic surgery reveals a high percentage of asymptomatic infections. *J. Bone Joint Surg. Am.* **102**, 1379–1388 (2020).
230. C.-X. Guo *et al*., Epidemiological and clinical features of pediatric COVID-19. *BMC Med.* **18**, 250 (2020).
231. T. Han, Outbreak investigation: Transmission of COVID-19 started from a spa facility in a local community in Korea. *Epidemiol. Health* **42**, e2020056 (2020).

232. T. Han *et al*., The epidemiological characteristics of cluster transmission of coronavirus disease 2019 (COVID-19): A multi-center study in Jiangsu Province. *Am. J. Transl. Res.* **12**, 6434–6444 (2020).
233. M. He *et al*., Epidemiological and clinical characteristics of 35 children with COVID-19 in Beijing, China. *Pediatr. Investig.* **4**, 230–235 (2020).
234. L. Hempel *et al*., SARS-CoV-2 infections in cancer outpatients-most infected patients are asymptomatic carriers without impact on chemotherapy. *Cancer Med.* **9**, 8020–8028 (2020).
235. I. Herraiz *et al*., Universal screening for SARS-CoV-2 before labor admission during Covid-19 pandemic in Madrid. *J. Perinat. Med.* **48**, 981–984 (2020).
236. D. Hijnen *et al*., SARS-CoV-2 transmission from presymptomatic meeting attendee, Germany. *Emerg. Infect. Dis.* **26**, 1935–1937 (2020).
237. J.-M. Hong *et al*., Epidemiological characteristics and clinical features of patients infected with the COVID-19 virus in Nanchang, Jiangxi, China. *Front. Med. (Lausanne)* **7**, 571069 (2020).
238. X. Hu *et al*., Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) vertical transmission in neonates born to mothers with Coronavirus Disease 2019 (COVID-19) pneumonia. *Obstet. Gynecol.* **136**, 65–67 (2020).
239. A. J. Ing, C. Cocks, J. P. Green, COVID-19: In the footsteps of Ernest Shackleton. *Thorax* **75**, 693–694 (2020).
240. S. Inui *et al*., Chest CT findings in cases from the cruise ship *Diamond Princess* with Coronavirus Disease (COVID-19). *Radiol. Cardiothorac. Imaging* **2**, e200110 (2020).
241. L. P. Jatt *et al*., Widespread severe acute respiratory coronavirus virus 2 (SARS-CoV-2) laboratory surveillance program to minimize asymptomatic transmission in high-risk inpatient and congregate living settings. *Infect. Control Hosp. Epidemiol.* **41**, 1331–1334 (2020).
242. J.-Y. Liu, T.-J. Chen, S.-J. Hwang, Analysis of community-acquired COVID-19 cases in Taiwan. *J. Chin. Med. Assoc.* **83**, 1087–1092 (2020).
243. C.-Y. Jung *et al*., Clinical characteristics of asymptomatic patients with COVID-19: A nationwide cohort study in South Korea. *Int. J. Infect. Dis.* **99**, 266–268 (2020).
244. J. Just, M.-T. Puth, F. Regenold, K. Weckbecker, M. Bleckwenn, Risk factors for a positive SARS-CoV-2 PCR in patients with common cold symptoms in a primary care setting - A retrospective analysis based on a joint documentation standard. *BMC Fam. Pract.* **21**, 251 (2020).
245. Expert Taskforce for the COVID-19 Cruise Ship Outbreak, Epidemiology of COVID-19 outbreak on cruise ship quarantined at Yokohama, Japan, February 2020. *Emerg. Infect. Dis.* **26**, 2591–2597 (2020).
246. A. M. Kassem *et al*., SARS-CoV-2 infection among healthcare workers of a gastroenterological service in a tertiary care facility. *Arab J. Gastroenterol.* **21**, 151–155 (2020).
247. K. Bai *et al*., Clinical analysis of 25 COVID-19 infections in children. *Pediatr. Infect. Dis. J.* **39**, e100–e103 (2020).
248. A. Khalil, R. Hill, S. Ladhani, K. Pattisson, P. O'Brien, Severe acute respiratory syndrome coronavirus 2 in pregnancy: Symptomatic pregnant women are only the tip of the iceberg. *Am. J. Obstet. Gynecol.* **223**, 296–297 (2020).
249. M. Ki, Task Force for 2019-nCoV, Epidemiologic characteristics of early cases with 2019 novel coronavirus (2019-nCoV) disease in Korea. *Epidemiol. Health* **42**, e2020007 (2020).
250. N. Koizumi, A. B. Siddique, A. Andalibi, Assessment of SARS-CoV-2 transmission among attendees of live concert events in Japan using contact-tracing data. *J. Travel Med.* **27**, taaa096 (2020).
251. COVID-19 National Emergency Response Center, Epidemiology and Case Management Team, Korea Centers for Disease Control and Prevention, Early epidemiological and clinical characteristics of 28 cases of coronavirus disease in South Korea. *Osong Public Health Res. Perspect.* **11**, 8–14 (2020).
252. V. B. Kute *et al*., Clinical profile and outcome of COVID-19 in 250 kidney transplant recipients: A multicenter cohort study from India. *Transplantation* **105**, 851–860 (2021).
253. S. M. LaCourse *et al*., Low prevalence of SARS-CoV-2 among pregnant and postpartum patients with universal screening in Seattle, Washington. *Clin. Infect. Dis.* **72**, 869–872 (2021).
254. X. Lai *et al*., Coronavirus Disease 2019 (COVID-2019) infection among health care workers and implications for prevention measures in a Tertiary Hospital in Wuhan, China. *JAMA Netw. Open* **3**, e209666 (2020).
255. E. Lavezzo *et al*., Suppression of a SARS-CoV-2 outbreak in the Italian municipality of Vo'. *Nature* **584**, 425–429 (2020).
256. A. G. Letizia *et al*., SARS-CoV-2 transmission among marine recruits during quarantine. *N. Engl. J. Med.* **383**, 2407–2416 (2020).
257. S. S. Lewis *et al*., Early experience with universal pre-procedural testing for SARS-CoV-2 in a relatively low-prevalence area. *Infect. Control Hosp. Epidemiol.* **42**, 341–343 (2021).
258. Y. Li *et al*., Asymptomatic and symptomatic patients with non-severe Coronavirus Disease (COVID-19) have similar clinical features and virological courses: A retrospective single center study. *Front. Microbiol.* **11**, 1570 (2020).
259. J. Li *et al*., Aggressive quarantine measures reduce the high morbidity of COVID-19 in patients on maintenance hemodialysis and medical staff of hemodialysis facilities in Wuhan, China. *Kidney Dis.* **6**, 271–283 (2020).
260. X. Li *et al*., Differences in clinical features and laboratory results between adults and children with SARS-CoV-2 infection. *BioMed Res. Int.* **2020**, 6342598 (2020).
261. S. Lin *et al*., Epidemiological and clinical characteristics of 161 discharged cases with coronavirus disease 2019 in Shanghai, China. *BMC Infect. Dis.* **20**, 780 (2020).
262. J.-Y. Liu, T.-J. Chen, S.-J. Hwang, Analysis of imported cases of COVID-19 in Taiwan: A nationwide study. *Int. J. Environ. Res. Public Health* **17**, 3311 (2020).
263. Epidemiology Working Group for NCIP Epidemic Response, Chinese Center for Disease Control and Prevention, The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China. *Zhonghua Liu Xing Bing Xue Za Zhi* **41**, 145–151 (2020).
264. L. Liu *et al*., Optimizing screening strategies for coronavirus disease 2019: A study from Middle China. *J. Infect. Public Health* **13**, 868–872 (2020).
265. S. Liu *et al*., Clinical characteristics and risk factors of patients with severe COVID-19 in Jiangsu province, China: A retrospective multicentre cohort study. *BMC Infect. Dis.* **20**, 584 (2020).
266. J. Liu, J. Huang, D. Xiang, Large SARS-CoV-2 outbreak caused by asymptomatic traveler, China. *Emerg. Infect. Dis.* **26**, 2260–2263 (2020).
267. P. Liu *et al*., Characteristics and effectiveness of the Coronavirus disease 2019 (COVID-19) prevention and control in a representative city in China. *Med. Sci. Monit.* **26**, e927472 (2020).
268. E. M. Lokken *et al*., Clinical characteristics of 46 pregnant women with a severe acute respiratory syndrome coronavirus 2 infection in Washington State. *Am. J. Obstet. Gynecol.* **223**, 911.e1–911.e14 (2020).
269. V. London *et al*., The relationship between status at presentation and outcomes among pregnant women with COVID-19. *Am. J. Perinatol.* **37**, 991–994 (2020).
270. A. S. Lopez *et al*., Transmission dynamics of COVID-19 outbreaks associated with child care facilities—Salt Lake City, Utah, April-July 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 1319–1323 (2020).
271. X. Lu *et al*., SARS-CoV-2 infection in children. *N. Engl. J. Med.* **382**, 1663–1665 (2020).
272. Field briefing: Diamond Princess COVID-19 cases, 20 Feb update. https://www.niid.go.jp/niid/en/2019-ncov-e/9417-covid-dp-fe-o2.html. Accessed 29 January 2021.
273. T. D. A. Ly *et al*., Pattern of SARS-CoV-2 infection among dependant elderly residents living in long-term care facilities in Marseille, France, March-June 2020. *Int. J. Antimicrob. Agents* **56**, 106219 (2020).
274. F. Maechler *et al*., Epidemiological and clinical characteristics of SARS-CoV-2 infections at a testing site in Berlin, Germany, March and April 2020-a cross-sectional study. *Clin. Microbiol. Infect.* **26**, 1685.e7–1685.e12 (2020).
275. F. Maggiolo *et al*., SARS-CoV-2 infection in persons living with HIV: A single center prospective cohort. *J. Med. Virol.* **93**, 1145–1149 (2021).
276. R. Malheiro *et al*., Effectiveness of contact tracing and quarantine on reducing COVID-19 transmission: A retrospective cohort study. *Public Health* **189**, 54–59 (2020).
277. H. C. Maltezou *et al*., Transmission dynamics of SARS-CoV-2 within families with children in Greece: A study of 23 clusters. *J. Med. Virol.* **93**, 1414–1420 (2021).
278. H. C. Maltezou *et al*., Children and adolescents with SARS-CoV-2 infection: Epidemiology, clinical course and viral loads. *Pediatr. Infect. Dis. J.* **39**, e388–e392 (2020).
279. S. Mao *et al*., Epidemiological analysis of 67 local COVID-19 clusters in Sichuan Province, China. *BMC Public Health* **20**, 1525 (2020).
280. A. Marossy *et al*., A study of universal SARS-CoV-2 RNA testing of residents and staff in a large group of care homes in South London. *J. Infect. Dis.* **223**, 381–388 (2021).
281. C. Martin *et al*., Dynamics of SARS-CoV-2 RT-PCR positivity and seroprevalence among high-risk healthcare workers and hospital staff. *J. Hosp. Infect.* **106**, 102–106 (2020).
282. F. Martini *et al*., On cancer, COVID-19, and CT scans: A monocentric retrospective study. *J. Clin. Med.* **9**, 3935 (2020).
283. C. Massarotti *et al*., Asymptomatic SARS-CoV-2 infections in pregnant patients in an Italian city during the complete lockdown. *J. Med. Virol.* **93**, 1758–1760 (2021).
284. M. K. Meena, M. Singh, P. K. Panda, M. K. Bairwa, Non-COVID area of a tertiary care hospital: A major source of nosocomial COVID-19 transmission. *J. Family Community Med.* **27**, 212–215 (2020).
285. M. Melgosa *et al*., SARS-CoV-2 infection in Spanish children with chronic kidney pathologies. *Pediatr. Nephrol.* **35**, 1521–1524 (2020).
286. N. Menachemi *et al*., Population point prevalence of SARS-CoV-2 infection based on a statewide random sample—Indiana, April 25-29, 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 960–964 (2020).
287. T. Menting *et al*., Low-threshold SARS-CoV-2 testing facility for hospital staff: Prevention of COVID-19 outbreaks? *Int. J. Hyg. Environ. Health* **231**, 113653 (2021).
288. E. S. Miller, W. A. Grobman, A. Sakowicz, J. Rosati, A. M. Peaceman, Clinical implications of universal Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) testing in pregnancy. *Obstet. Gynecol.* **136**, 232–234 (2020).
289. A. Mostafa *et al*., Universal COVID-19 screening of 4040 health care workers in a resource-limited setting: An Egyptian pilot model in a university with 12 public hospitals and medical centers. *Int. J. Epidemiol.* **50**, 50–61 (2021).
290. E. Namal, Single center experience on screening oncology patients for covid-19 before anti-cancer treatment. *Int. J. Hematol. Oncol.* **30**, 207–212 (2020).
291. H. Nishiura *et al*., The rate of underascertainment of novel coronavirus (2019-nCoV) infection: Estimation using Japanese passengers data on evacuation flights. *J. Clin. Med.* **9**, 419 (2020).
292. L. Norsa *et al*., Asymptomatic severe acute respiratory syndrome coronavirus 2 infection in patients with inflammatory bowel disease under biologic treatment. *Gastroenterology* **159**, 2229–2231.e2 (2020).
293. S. Ornaghi *et al*., Performance of an extended triage questionnaire to detect suspected cases of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) infection in obstetric patients: Experience from two large teaching hospitals in Lombardy, Northern Italy. *PLoS One* **15**, e0239173 (2020).
294. S. Ossareh, M. Bagheri, M. Abbasi, S. Abolfathi, A. Bohlooli, Role of screening for COVID-19 in hemodialysis wards, results of a single center study. *Iran. J. Kidney Dis.* **14**, 389–398 (2020).
295. Y. Oster *et al*., Proactive screening approach for SARS-CoV-2 among healthcare workers. *Clin. Microbiol. Infect.* **27**, 155–156 (2021).
296. L. Panagiotakopoulos *et al*., SARS-CoV-2 infection among hospitalized pregnant women: Reasons for admission and pregnancy characteristics—Eight U.S. Health Care Centers, March 1-May 30, 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 1355–1359 (2020).

297. N. Parri *et al*., Characteristic of COVID-19 infection in pediatric patients: Early findings from two Italian pediatric research networks. *Eur. J. Pediatr.* **179**, 1315–1323 (2020).
298. N. Parri *et al*., COVID-19 in 17 Italian pediatric emergency departments. *Pediatrics* **146**, e20201235 (2020).
299. E. T. Patberg *et al*., Coronavirus disease 2019 infection and placental histopathology in women delivering at term. *Am. J. Obstet. Gynecol.* **224**, 382.e1–382.e18 (2021).
300. A. B. Patel *et al*., SARS-CoV-2 point prevalence among asymptomatic hospitalized children and subsequent healthcare worker evaluation. *J. Pediatric Infect. Dis. Soc.* **9**, 617–619 (2020).
301. A. Pavli *et al*., In-flight transmission of COVID-19 on flights to Greece: An epidemiological analysis. *Travel Med. Infect. Dis.* **38**, 101882 (2020).
302. D. C. Payne *et al*., SARS-CoV-2 infections and serologic responses from a sample of U.S. Navy Service Members—USS Theodore Roosevelt, April 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 714–721 (2020).
303. I. Petersen, A. Phillips, Three quarters of people with SARS-CoV-2 infection are asymptomatic: Analysis of English household survey data. *Clin. Epidemiol.* **12**, 1039–1043 (2020).
304. J.-P. Pirnay *et al*., Study of a SARS-CoV-2 outbreak in a Belgian military education and training center in Maradi, Niger. *Viruses* **12**, 949 (2020).
305. M. Pollán *et al*., Prevalence of SARS-CoV-2 in Spain (ENE-COVID): A nationwide, population-based seroepidemiological study. *Lancet* **396**, 535–544 (2020).
306. M. Prabhu *et al*., Pregnancy and postpartum outcomes in a universally tested population for SARS-CoV-2 in New York City: A prospective cohort study. *BJOG* **127**, 1548–1556 (2020).
307. I. W. Pray *et al*., COVID-19 outbreak at an overnight summer school retreat—Wisconsin, July-August 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 1600–1604 (2020).
308. P. Xingqiang *et al*., Study on transmission dynamic of 15 clusters of coronavirus disease 2019 cases in Ningbo. *Zhonghua Liu Xing Bing Xue Za Zhi* **41**, E066 (2020).
309. H. Qiu *et al*., Clinical and epidemiological features of 36 children with coronavirus disease 2019 (COVID-19) in Zhejiang, China: An observational cohort study. *Lancet Infect. Dis.* **20**, 689–696 (2020).
310. J. H. Rogers *et al*., Characteristics of COVID-19 in homeless shelters: A community-based surveillance study. *Ann. Intern. Med.* **174**, 42–49 (2021).
311. A. C. Roxby *et al*., Outbreak investigation of COVID-19 among residents and staff of an independent and assisted living community for older adults in Seattle, Washington. *JAMA Intern. Med.* **180**, 1101–1105 (2020).
312. G. Sacco, G. Foucault, O. Briere, C. Annweiler, COVID-19 in seniors: Findings and lessons from mass screening in a nursing home. *Maturitas* **141**, 46–52 (2020).
313. I. H. Huerta Saenz, J. C. Elías Estrada, K. Campos Del Castillo, R. Muñoz Taya, J. C. Coronado, Características materno perinatales de gestantes COVID-19 en un hospital nacional de Lima, Perú [in Spanish]. *Rev. Peru. Ginecol. Obstet.*, 10.31403/rpgo.v66i2245 (2020).
314. A. Sakowicz *et al*., Risk factors for severe acute respiratory syndrome coronavirus 2 infection in pregnant women. *Am. J. Obstet. Gynecol. MFM* **2**, 100198 (2020).
315. M. Saluja, D. Pillai, S. Jeliya, N. Bauddh, R. Chandel, COVID 19-clinical profile, radiological presentation, prognostic predictors, complications and outcome: A perspective from the Indian subcontinent. *J. Assoc. Physicians India* **68**, 13–18 (2020).
316. S. M. Samrah *et al*., COVID-19 outbreak in Jordan: Epidemiological features, clinical characteristics, and laboratory findings. *Ann. Med. Surg. (Lond.)* **57**, 103–108 (2020).
317. G. V. Sanchez *et al*., Initial and repeated point prevalence surveys to inform SARS-CoV-2 infection prevention in 26 skilled nursing facilities—Detroit, Michigan, March-May 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 882–886 (2020).
318. S. R. Sastry *et al*., Universal screening for the SARS-CoV-2 virus on hospital admission in an area with low COVID-19 prevalence. *Infect. Control Hosp. Epidemiol.* **41**, 1231–1233 (2020).
319. S. Saurabh *et al*., Dynamics of SARS-CoV-2 transmission among Indian nationals evacuated from Iran. *Disaster Med. Public Health Prep.*, 10.1017/dmp.2020.393 (2020).
320. R. Savirón-Cornudella *et al*., Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) universal screening in gravids during labor and delivery. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **256**, 400–404 (2021).
321. Q. Shen *et al*., Novel coronavirus infection in children outside of Wuhan, China. *Pediatr. Pulmonol.* **55**, 1424–1429 (2020).
322. A. K. Sharma *et al*., Epidemiological and clinical profile of COVID-19 in Nepali children: An initial experience. *J. Nepal Paediatr. Soc.* **40**, 202–209 (2020).
323. L. Shi *et al*., Quarantine at home may not enough!-from the epidemiological data in Shaanxi Province of China. *BMC Res. Notes* **13**, 506 (2020).
324. R. G. Shmakov *et al*., Clinical course of novel COVID-19 infection in pregnant women. *J. Matern. Fetal Neonatal Med.*, 10.1080/14767058.2020.1850683 (2020).
325. J. S. Singer *et al*., Low prevalence (0.13%) of COVID-19 infection in asymptomatic pre-operative/pre-procedure patients at a large, academic medical center informs approaches to perioperative care. *Surgery* **168**, 980–986 (2020).
326. N. Sugano, W. Ando, W. Fukushima, Cluster of severe acute respiratory syndrome coronavirus 2 infections linked to music clubs in Osaka, Japan. *J. Infect. Dis.* **222**, 1635–1640 (2020).
327. D. Sun *et al*., Children infected with SARS-CoV-2 from family clusters. *Front. Pediatr.* **8**, 386 (2020).
328. D. Sutton, K. Fuchs, M. D'Alton, D. Goffman, Universal screening for SARS-CoV-2 in women admitted for delivery. *N. Engl. J. Med.* **382**, 2163–2164 (2020).
329. M. P. Tambe *et al*., An epidemiological study of laboratory confirmed COVID-19 cases admitted in a tertiary care hospital of Pune, Maharashtra. *Indian J. Public Health* **64**, S183–S187 (2020).
330. T. Xin *et al*., Clinical characteristics analysis of 13 cases of novel coronavirus infection in children in Changsha. *Chinese Journal of Contemporary Pediatrics* **22**, 294–298 (2020).
331. O. Tang *et al*., Outcomes of nursing home COVID-19 patients by initial symptoms and comorbidity: Results of universal testing of 1970 residents. *J. Am. Med. Dir. Assoc.* **21**, 1767–1773.e1 (2020).
332. T. Ali *et al*., Coronavirus disease-19: Disease severity and outcomes of solid organ transplant recipients: Different spectrum of disease in different populations? *Transplantation* **105**, 121–127 (2021).
333. J. W. Thompson Jr. *et al*., An epidemiologic study of COVID-19 patients in a state psychiatric hospital: High penetrance with early CDC guidelines. *Psychiatr. Serv.* **71**, 1285–1287 (2020).
334. S. Tian *et al*., Characteristics of COVID-19 infection in Beijing. *J. Infect.* **80**, 401–406 (2020).
335. M.-J. Trahan, C. Mitric, I. Malhamé, H. A. Abenhaim, Screening and testing pregnant patients for SARS-CoV-2: First-wave experience of a designated COVID-19 hospitalization centre in Montreal. *J. Obstet. Gynaecol. Can.* **43**, 571–575 (2021).
336. T. Venkataram *et al*., Deployment of neurosurgeons at the warfront against coronavirus disease of 2019 (COVID-19). *World Neurosurg.* **144**, e561–e567 (2020).
337. M. Wang *et al*., Epidemiological characteristics and transmission dynamics of paediatric cases with coronavirus disease 2019 in Hubei province, China. *J. Paediatr. Child Health* **57**, 637–645 (2021).
338. S. Wanwan *et al*., Epidemiological characteristics of 2019 novel coronavirus family clustering in Zhejiang Province. *Chin. J. Prev. Med* **54**, E027 (2020).
339. X. Xu *et al*., The cumulative rate of SARS-CoV-2 infection in Chinese hemodialysis patients. *Kidney Int. Rep.* **5**, 1416–1421 (2020).
340. J. Yan *et al*., Coronavirus disease 2019 in pregnant women: A report based on 116 cases. *Am. J. Obstet. Gynecol.* **223**, 111.e1–111.e14 (2020).
341. J. Yang *et al*., Clinical characteristics and eosinophils in young SARS-CoV-2-positive Chinese travelers returning to Shanghai. *Front. Public Health* **8**, 368 (2020).
342. M. Yassa *et al*., Outcomes of universal SARS-CoV-2 testing program in pregnant women admitted to hospital and the adjuvant role of lung ultrasound in screening: A prospective cohort study. *J. Matern. Fetal Neonatal Med.* **33**, 3820–3826 (2020).
343. B. C. Cura Yayla *et al*., Characteristics and management of children with COVID-19 in Turkey. *Balkan Med. J.* **37**, 341–347 (2020).
344. B. C. C. Yayla, K. Aykac, Y. Ozsurekci, M. Ceyhan, Characteristics and management of children with COVID-19 in a tertiary care hospital in Turkey. *Clin. Pediatr.* **60**, 170–177 (2021).
345. K. Yılmaz *et al*., Evaluation of the novel coronavirus disease in Turkish children: Preliminary outcomes. *Pediatr. Pulmonol.* **55**, 3587–3594 (2020).
346. L. X. Ye *et al*., Investigation of a cluster epidemic of COVID-19 in Ningbo. *Zhonghua Liu Xing Bing Xue Za Zhi* **41**, 2029–2033 (2020).
347. J. C. Yombi, J. De Greef, P. Bernard, L. Belkhir, Testing of patients and coronavirus disease 2019 (COVID-19) infection before scheduled deliveries. *J. Perinat. Med.* **48**, 995–996 (2020).
348. H. Yue *et al*., Clinical characteristics of coronavirus disease 2019 in Gansu province, China. *Ann. Palliat. Med.* **9**, 1404–1412 (2020).
349. T. Zhan *et al*., Retrospective analysis of clinical characteristics of 405 patients with COVID-19. *J. Int. Med. Res.* **48**, 300060520949039 (2020).
350. M. 'A. I. A. Zamzuri *et al*., Epidemiological characteristics of COVID-19 in Seremban, Negeri Sembilan, Malaysia. *Open Access Maced. J. Med. Sci.* **8**, 471–475 (2020).
351. L. Zhang, S. Huang, Clinical features of 33 cases in children infected with SARS-CoV-2 in Anhui Province, China—A multi-center retrospective cohort study. *Front. Public Health* **8**, 255 (2020).
352. S. Zhang *et al*., Factors associated with asymptomatic infection in health-care workers with severe acute respiratory syndrome coronavirus 2 infection in Wuhan, China: A multicentre retrospective cohort study. *Clin. Microbiol. Infect.* **26**, 1670–1675 (2020).
353. Z. Q. Deng *et al*., Analysis on transmission chain of a cluster epidemic of COVID-19, Nanchang. *Zhonghua Liu Xing Bing Xue Za Zhi* **41**, 1420–1423 (2020).
354. L. Zou *et al*., SARS-CoV-2 viral load in upper respiratory specimens of infected patients. *N. Engl. J. Med.* **382**, 1177–1179 (2020).
355. R. H. El-Sokkary *et al*., Characteristics and predicting factors of Corona Virus Disease-2019 (COVID-19) among healthcare providers in a developing country. *PLoS One* **16**, e0245672 (2021).
356. M. F. Kristiansen *et al*., Epidemiology and clinical course of first wave coronavirus disease cases, Faroe Islands. *Emerg. Infect. Dis.* **27**, 749–758 (2021).
357. S. C. Reale *et al*., Patient characteristics associated with SARS-CoV-2 infection in parturients admitted for labour and delivery in Massachusetts during the spring 2020 surge: A prospective cohort study. *Paediatr. Perinat. Epidemiol.* **35**, 24–33 (2021).
358. Y. Min *et al*., Clinical characteristics of deceased hemodialysis patients affected by COVID-19. *Int. Urol. Nephrol.* **53**, 797–802 (2021).
359. J. A. Huete-Pérez *et al*., First report on prevalence of SARS-CoV-2 infection among health-care workers in Nicaragua. *PLoS One* **16**, e0246084 (2021).
360. F. W. Arnold, S. Bishop, L. Oppy, L. Scott, G. Stevenson, Surveillance testing reveals a significant proportion of hospitalized patients with SARS-CoV-2 are asymptomatic. *Am. J. Infect. Control* **49**, 281–285 (2021).
361. B. F. Bigelow *et al*., Outcomes of universal COVID-19 testing following detection of incident cases in 11 long-term care facilities. *JAMA Intern. Med.* **181**, 127–129 (2021).
362. C. Bayle *et al*., Asymptomatic SARS COV-2 carriers among nursing home staff: A source of contamination for residents? *Infect. Dis. Now* **51**, 197–200 (2021).
363. L. Wang *et al*., Source investigation on a familiar cluster of coronavirus disease 2019 in Dandong city of Liaoning Province. *Zhonghua Yu Fang Yi Xue Za Zhi* **55**, 120–122 (2021).

364. T. Scheier *et al*., Universal admission screening for SARS-CoV-2 infections among hospitalized patients, Switzerland, 2020. *Emerg. Infect. Dis.* **27**, 404–410 (2021).
365. H. H. Adetola *et al*., Clinical presentations and management of COVID-19 infected children seen in a district health facility in Kambia, northern Sierra Leone. *Pan Afr. Med. J.* **37**, 28 (2020).
366. E. Camara *et al*., Epidemiological and clinical profile of children with Coronavirus disease (COVID-19) at the Center for the Treatment of Epidemics and Infection Prevention (CTEIP) of the University Hospital of Donka in Conakry. *Pan Afr. Med. J.* **37**, 363 (2020).
367. A. Jones *et al*., Assessment of day-7 postexposure testing of asymptomatic contacts of COVID-19 patients to evaluate early release from quarantine—Vermont, May-November 2020. *MMWR Morb. Mortal. Wkly. Rep.* **70**, 12–13 (2021).
368. T. Khondaker *et al*., Clinical profile and outcome of COVID -19 in children with pre-existing renal disease. *Int. J. Pediatr. Nephrol.* **9**, 1–6 (2021).
369. N. Marcus *et al*., Minor clinical impact of COVID-19 pandemic on patients with primary immunodeficiency in Israel. *Front. Immunol.* **11**, 614086 (2021).
370. C. Martin-Villares, M. Bernal-Sprekelsen, C. P. Molina-Ramirez, M. Bartolome-Benito, COVID ORL ESP Collaborative Group, Risk of contagion of SARS-CoV-2 among oto-rhinolaryngologists in Spain during the "Two waves." *Eur. Arch. Otorhinolaryngol.*, 10.1007/s00405-020-06582-8 (2021).
371. M. A. Almadhi *et al*., The high prevalence of asymptomatic SARS-CoV-2 infection reveals the silent spread of COVID-19. *Int. J. Infect. Dis.* **105**, 656–661 (2021).
372. H. G. Atakla *et al*., COVID-19 infection in pediatric subjects: Study of 36 cases in Conakry. *Pan Afr. Med. J.* **37**, 42 (2020).
373. J. K. Bender, M. Brandl, M. Höhle, U. Buchholz, N. Zeitlmann, Analysis of asymptomatic and presymptomatic transmission in SARS-CoV-2 outbreak, Germany, 2020. *Emerg. Infect. Dis.* **27**, 1159–1163 (2021).
374. M. Berry *et al*., Clinical stratification of pregnant COVID-19 patients based on severity: A single academic center experience. *Am. J. Perinatol.* **38**, 515–522 (2021).
375. M. N. S. Cabraal, R. I. U. Samarawickrama, K. A. R. R. Kodithuwakku, S. D. Viswakula, S. R. Lantra, Nationwide descriptive study of COVID-19 in children below the age of 14 years in Sri Lanka. *Sri Lanka J. Child Health* **50**, 103 (2021).
376. C. Cesilia, S. Sudarmaji, D. Setiabudi, H. M. Nataprawira, Case report of a COVID-19 family cluster originating from a boarding school. *Paediatr. Indones.* **61**, 53–60 (2021).
377. L. Dbeibo *et al*., Assessment of a universal preprocedural screening program for COVID-19. *Infect. Control Hosp. Epidemiol.*, 10.1017/ice.2021.40 (2021).
378. B. E. Dixon *et al*., Symptoms and symptom clusters associated with SARS-CoV-2 infection in community-based populations: Results from a statewide epidemiological study. *PLoS One* **16**, e0241875 (2021).
379. H. Fakhim *et al*., Asymptomatic carriers of coronavirus disease 2019 among healthcare workers in Isfahan, Iran. *Future Virol.* **16**, 93–98 (2021).
380. S. Hu *et al*., Infectivity, susceptibility, and risk factors associated with SARS-CoV-2 transmission under intensive contact tracing in Hunan, China. *Nat. Commun.* **12**, 1533 (2021).
381. S. Isoldi *et al*., The comprehensive clinic, laboratory, and instrumental evaluation of children with COVID-19: A 6-months prospective study. *J. Med. Virol.* **93**, 3122–3132 (2021).
382. S. Jani *et al*., Clinical characteristics of mother-infant dyad and placental pathology in COVID-19 cases in predominantly African American population. *AJP Rep.* **11**, e15–e20 (2021).
383. V. B. Kute *et al*., A multicenter cohort study from India of 75 kidney transplants in recipients recovered after COVID-19. *Transplantation* **105**, 1423–1432 (2021).
384. J. E. Marcus *et al*., Risk factors associated with COVID-19 transmission among US air force trainees in a congregant setting. *JAMA Netw. Open* **4**, e210202 (2021).
385. R. Miyahara *et al*., Familial clusters of coronavirus disease in 10 prefectures, Japan, February-May 2020. *Emerg. Infect. Dis.* **27**, 915–918 (2021).
386. P. S. Myles *et al*., COVID-19 risk in elective surgery during a second wave: A prospective cohort study. *ANZ J. Surg.* **91**, 22–26 (2021).
387. K. Nakajo, H. Nishiura, Transmissibility of asymptomatic COVID-19: Data from Japanese clusters. *Int. J. Infect. Dis.* **105**, 236–238 (2021).
388. M. de Miguel Negro *et al*., Pre-operative prevalence of asymptomatic carriers of COVID-19 in hospitals in Catalonia during the first wave after the resumption of surgical activity. *Cir. Esp. (Engl. Ed.)*, 10.1016/j.ciresp.2021.01.014 (2021).
389. E. Oduro-Mensah *et al*., Clinical features of COVID-19 in Ghana: Symptomatology, illness severity and comorbid non-communicable diseases. *Ghana Med. J.* **5**, 23–32 (2020).
390. R. V. Randremanana *et al*., The COVID-19 epidemic in Madagascar: Clinical description and laboratory results of the first wave, March-September 2020. *Influenza Other Respi. Viruses* **15**, 457–468 (2021).
391. G. Sabetian *et al*., COVID-19 infection among healthcare workers: A cross-sectional study in southwest Iran. *Virol. J.* **18**, 58 (2021).
392. R. Savirón-Cornudella *et al*., Screening of severe acute respiratory syndrome coronavirus-2 infection during labor and delivery using polymerase chain reaction and immunoglobulin testing. *Life Sci.* **271**, 119200 (2021).
393. V. Singh, A. Choudhary, M. R. Datta, A. Ray, Maternal and neonatal outcomes of COVID-19 in pregnancy: A single-centre observational study. *Cureus* **13**, e13184 (2021).
394. N. Sharma, N. Seehra, S. Kabra, Pregnancy with covid-19 infection and fetomaternal outcomes. *J. Evol. Med. Dent. Sci.* **10**, 23–27 (2021).
395. A. Soriano-Arandes *et al*., Household SARS-CoV-2 transmission and children: A network prospective study. *Clin. Infect. Dis.*, 10.1093/cid/ciab228 (2021).
396. R. N. Stadler *et al*., Systematic screening on admission for SARS-CoV-2 to detect asymptomatic infections. *Antimicrob. Resist. Infect. Control* **10**, 44 (2021).
397. S. Tawar, G. Diva Reddy, S. Ray, N. Chawla, S. Garg, Rapid response and mitigation measures in control of COVID-19 cases in an industrial warehouse of Western Maharashtra, India. *J. Mar. Med. Soc.* **22**, 220 (2020).
398. L. K. Tompkins *et al*., Mass SARS-CoV-2 testing in a dormitory-style correctional facility in Arkansas. *Am. J. Public Health* **111**, 907–916 (2021).
399. J. Villa *et al*., Results of preoperative screening for COVID-19 correlate with the incidence of infection in the general population—A tertiary care experience. *Hosp. Pract.* **49**, 216–220 (2021).
400. W. Xie *et al*., Infection and disease spectrum in individuals with household exposure to SARS-CoV-2: A family cluster cohort study. *J. Med. Virol.* **93**, 3033–3046 (2021).
401. N. Van Vinh Chau *et al*., The natural history and transmission potential of asymptomatic SARS-CoV-2 infection. *Clin. Infect. Dis.* **71**, 2679–2687 (2020).
402. H. Xu *et al*., A follow-up study of children infected with SARS-CoV-2 from western China. *Ann. Transl. Med.* **8**, 10 (2020).
403. H. P. Wu, Clinical features of coronavirus disease 2019 in children aged <18 years in Jiangxi, China: An analysis of 23 cases. *Chinese Journal of Contemporary Pediatrics* **22**, 419–424 (2020).
404. M. Yaoling, Analysis of the clinical characteristics of 115 children with novel coronavirus infection. *Chinese Journal of Contemporary Pediatrics* **22**, 290 (2020).
405. Z. Yun, Clinical features and chest CT findings of 2019 coronavirus disease in infants and young children. *Chinese Journal of Contemporary Pediatrics* **22**, 215 (2020).
406. N. van Doremalen *et al*., Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *N. Engl. J. Med.* **382**, 1564–1567 (2020).
407. P. K. C. Cheng *et al*., Viral shedding patterns of coronavirus in patients with probable severe acute respiratory syndrome. *Lancet* **363**, 1699–1700 (2004).
408. D. P. Oran, E. J. Topol, Prevalence of asymptomatic SARS-CoV-2 infection: A narrative review. *Ann. Intern. Med.* **173**, 362–367 (2020).
409. A. Kronbichler *et al*., Asymptomatic patients as a source of COVID-19 infections: A systematic review and meta-analysis. *Int. J. Infect. Dis.* **98**, 180–186 (2020).
410. The Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, The novel coronavirus pneumonia emergency response epidemiology team, the epidemiological characteristics of an outbreak of 2019 novel Coronavirus Diseases (COVID-19)—China, 2020. *China CDC Weekly* **2**, 113–122 (2020).
411. X. Zhou, Y. Li, T. Li, W. Zhang, Follow-up of asymptomatic patients with SARS-CoV-2 infection. *Clin. Microbiol. Infect.* **26**, 957–959 (2020).
412. N. H. L. Leung, C. Xu, D. K. M. Ip, B. J. Cowling, Review article: The fraction of influenza virus infections that are asymptomatic: A systematic review and meta-analysis. *Epidemiology* **26**, 862–872 (2015).
413. L. Furuya-Kanamori *et al*., Heterogeneous and dynamic prevalence of asymptomatic influenza virus infections. *Emerg. Infect. Dis.* **22**, 1052–1056 (2016).
414. A. Wilder-Smith *et al*., Asymptomatic SARS coronavirus infection among healthcare workers, Singapore. *Emerg. Infect. Dis.* **11**, 1142–1145 (2005).
415. F. Perrotta *et al*., COVID-19 and the elderly: Insights into pathogenesis and clinical decision-making. *Aging Clin. Exp. Res.* **32**, 1599–1608 (2020).
416. P. Mehta *et al*., COVID-19: Consider cytokine storm syndromes and immunosuppression. *Lancet* **395**, 1033–1034 (2020).
417. M. E. Carter-Timofte *et al*., Deciphering the role of host genetics in susceptibility to severe COVID-19. *Front. Immunol.* **11**, 1606 (2020).
418. P. Conti, A. Younes, Coronavirus COV-19/SARS-CoV-2 affects women less than men: Clinical response to viral infection. *J. Biol. Regul. Homeost. Agents* **34**, 339–343 (2020).
419. N. Le Bert *et al*., SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* **584**, 457–462 (2020).
420. D. Whyte *et al*., Mumps epidemiology in the mid-west of Ireland 2004-2008: Increasing disease burden in the university/college setting. *Euro Surveill.* **14**, 19182 (2009).
421. Centers for Disease Control and Prevention, Mumps outbreak on a university campus—California, 2011. *MMWR Morb. Mortal. Wkly. Rep.* **61**, 986–989 (2012).
422. J. Lessler *et al*., Outbreak of 2009 pandemic influenza A (H1N1) at a New York City school. *N. Engl. J. Med.* **361**, 2628–2636 (2009).
423. C. Stein-Zamir *et al*., A large COVID-19 outbreak in a high school 10 days after schools' reopening, Israel, May 2020. *Euro Surveill.* **25**, 2001352 (2020).
424. World Health Organization, Coronavirus disease (COVID-19). https://www.who.int/emergencies/diseases/novel-coronavirus-2019. Accessed 14 August 2020.
425. G. J. Griffith *et al*., Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nat. Commun.* **11**, 5749 (2020).
426. Centers for Disease Control and Prevention, Coronavirus Disease 2019 (COVID-19). https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html. Accessed 27 May 2020.
427. World Health Organization, Global research on coronavirus disease (COVID-19). https://www.who.int/emergencies/diseases/novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov. Accessed 4 May 2020.
428. J. A. Sterne *et al*., ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* **355**, i4919 (2016).
429. M. Borenstein, L. V. Hedges, J. P. T. Higgins, H. R. Rothstein, A basic introduction to fixed-effect and random-effects models for meta-analysis. *Res. Synth. Methods* **1**, 97–111 (2010).
430. R. DerSimonian, N. Laird, Meta-analysis in clinical trials. *Control. Clin. Trials* **7**, 177–188 (1986).
431. A. Agresti, B. A. Coull, Approximate is better than "exact" for interval estimation of binomial proportions. *null* **52**, 119–126 (1998).
432. J. Hartung, G. Knapp, On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Stat. Med.* **20**, 1771–1782 (2001).
433. M. Egger, G. Davey Smith, M. Schneider, C. Minder, Bias in meta-analysis detected by a simple, graphical test. *BMJ* **315**, 629–634 (1997).

Exhibit

41

JAMA Network Open

Original Investigation | Infectious Diseases

SARS-CoV-2 Transmission From People Without COVID-19 Symptoms

Michael A. Johansson, PhD; Talia M. Quandelacy, PhD, MPH; Sarah Kada, PhD; Pragati Venkata Prasad, MPH; Molly Steele, PhD, MPH; John T. Brooks, MD; Rachel B. Slayton, PhD, MPH; Matthew Biggerstaff, ScD, MPH; Jay C. Butler, MD

Abstract

IMPORTANCE Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the etiology of coronavirus disease 2019 (COVID-19), is readily transmitted person to person. Optimal control of COVID-19 depends on directing resources and health messaging to mitigation efforts that are most likely to prevent transmission, but the relative importance of such measures has been disputed.

OBJECTIVE To assess the proportion of SARS-CoV-2 transmissions in the community that likely occur from persons without symptoms.

DESIGN, SETTING, AND PARTICIPANTS This decision analytical model assessed the relative amount of transmission from presymptomatic, never symptomatic, and symptomatic individuals across a range of scenarios in which the proportion of transmission from people who never develop symptoms (ie, remain asymptomatic) and the infectious period were varied according to published best estimates. For all estimates, data from a meta-analysis was used to set the incubation period at a median of 5 days. The infectious period duration was maintained at 10 days, and peak infectiousness was varied between 3 and 7 days (−2 and +2 days relative to the median incubation period). The overall proportion of SARS-CoV-2 was varied between 0% and 70% to assess a wide range of possible proportions.

MAIN OUTCOMES AND MEASURES Level of transmission of SARS-CoV-2 from presymptomatic, never symptomatic, and symptomatic individuals.

RESULTS The baseline assumptions for the model were that peak infectiousness occurred at the median of symptom onset and that 30% of individuals with infection never develop symptoms and are 75% as infectious as those who do develop symptoms. Combined, these baseline assumptions imply that persons with infection who never develop symptoms may account for approximately 24% of all transmission. In this base case, 59% of all transmission came from asymptomatic transmission, comprising 35% from presymptomatic individuals and 24% from individuals who never develop symptoms. Under a broad range of values for each of these assumptions, at least 50% of new SARS-CoV-2 infections was estimated to have originated from exposure to individuals with infection but without symptoms.

CONCLUSIONS AND RELEVANCE In this decision analytical model of multiple scenarios of proportions of asymptomatic individuals with COVID-19 and infectious periods, transmission from asymptomatic individuals was estimated to account for more than half of all transmissions. In addition to identification and isolation of persons with symptomatic COVID-19, effective control of spread will require reducing the risk of transmission from people with infection who do not have symptoms. These findings suggest that measures such as wearing masks, hand hygiene, social distancing, and strategic testing of people who are not ill will be foundational to slowing the spread

Key Points

Question What proportion of coronavirus disease 2019 (COVID-19) spread is associated with transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) from persons with no symptoms?

Findings In this decision analytical model assessing multiple scenarios for the infectious period and the proportion of transmission from individuals who never have COVID-19 symptoms, transmission from asymptomatic individuals was estimated to account for more than half of all transmission.

Meaning The findings of this study suggest that the identification and isolation of persons with symptomatic COVID-19 alone will not control the ongoing spread of SARS-CoV-2.

+ Multimedia

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

Abstract (continued)

of COVID-19 until safe and effective vaccines are available and widely used.

JAMA Network Open. 2021;4(1):e2035057.
Corrected on February 12, 2021. doi:10.1001/jamanetworkopen.2020.35057

Introduction

As severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the novel coronavirus that causes coronavirus disease 2019 (COVID-19), began to spread globally, it became apparent that the virus, unlike the closely related SARS-CoV in the 2003 outbreak, could not be contained by symptom-based screening alone. Asymptomatic and clinically mild infections were uncommon during the 2003 SARS-CoV outbreak, and there were no reported instances of transmission from persons before the onset of symptoms.[1] SARS-CoV-2 spread faster than SARS-CoV, and accumulating evidence showed that SARS-CoV-2, unlike SARS-CoV, is transmitted from persons without symptoms. However, measures to reduce transmission from individuals who do not have COVID-19 symptoms have become controversial and politicized and have likely had negative effects on the economy and many societal activities. Optimal control of COVID-19 depends on directing resources and health messaging to mitigation efforts that are most likely to prevent transmission. The relative importance of mitigation measures that prevent transmission from persons without symptoms has been disputed. Determining the proportion of SARS-CoV-2 transmission that occurs from persons without symptoms is foundational to prioritizing control practices and policies.

Transmission by persons who are infected but do not have any symptoms can arise from 2 different infection states: presymptomatic individuals (who are infectious before developing symptoms) and individuals who never experience symptoms (asymptomatic infections, which we will refer to as never symptomatic). Early modeling studies of COVID-19 case data found that the generation interval of SARS-CoV-2 was shorter than the serial interval, indicating that the average time between 1 person being infected and that person infecting someone else was shorter than the average time between 1 person developing symptoms and the person they infected developing symptoms.[2-5] This finding meant that the epidemic was growing faster than would be expected if transmission were limited to the period of illness during which individuals were symptomatic. By the time a second generation of individuals was developing symptoms, a third generation was already being infected. Epidemiological data from early in the pandemic also suggested the possibility of presymptomatic transmission,[6,7] and laboratory studies confirmed that levels of viral RNA in respiratory secretions were already high at the time of symptom onset.[8-10]

Asymptomatic SARS-CoV-2 transmission also occurs because of individuals with infection who are never symptomatic (or who experience very mild or almost unrecognizable symptoms). The proportion of individuals with infection who never have apparent symptoms is difficult to quantify because it requires intensive prospective clinical sampling and symptom screening from a representative sample of individuals with and without infection. Nonetheless, evidence from household contact studies indicates that asymptomatic or very mild symptomatic infections occur,[11-14] and laboratory and epidemiological evidence suggests that individuals who never develop symptoms may be as likely as individuals with symptoms to transmit SARS-CoV-2 to others.[9,15,16]

Methods

The Centers for Disease Control and Prevention determined that this decision analytical study, which involved no enrollment of human subjects, did not require institutional review board approval. We used a simple model to assess the proportion of transmission from presymptomatic (ie, infectious before symptom onset), never symptomatic, and symptomatic individuals across a range of

scenarios in which we varied the timing of the infectious period to assess different contributions of presymptomatic transmission and the proportion of transmission from individuals who never develop symptoms (ie, remain asymptomatic).

For all estimates we used data from a meta-analysis of 8 studies from China to set the incubation period at a median of 5 days with 95% of symptomatic individuals developing symptoms by day 12.[17] Therefore the daily (t) probability of symptom onset (p_{so}) for individuals who develop symptoms was:

$$p_{so}(t) = F_{Log-Normal}(t,\text{logmean} = 1.63,\text{logsd} = 0.5).$$

To approximate a distribution of the infectious period, we made a baseline assumption that peak infectiousness occurs on average at the same time as the median incubation period, such that infectiousness begins prior to symptom onset (**Table**).[9,12,14-16,18,20] We then assumed that infectiousness (I) over time can be approximated by a γ density function and that the average person is infectious for as long as approximately 10 days (ie, 98% of transmission happens within a 10-day period)[11]:

$$I(t) = f_{\gamma}(t,\text{mode} = 5,\text{interval} = 10).$$

For all estimates, we maintained the infectious period duration as 10 days, but varied the mode between 3 and 7 days (−2 and +2 days relative to the median incubation period).

Uncertainty also remains about the proportion of individuals with infection who are never symptomatic (p_{ns}) and the relative contribution of these infections to transmission (r_{ns}). Estimates of p_{ns} range from single digits to more than 50%, many with potential biases related to the study population (eg, age, prevalence of comorbidities) and the extent of long-term follow-up[12-14,19,20] (Table). We made a baseline assumption that 30% of individuals with infection are never symptomatic and then assessed higher or lower assumptions. We also made a baseline assumption that individuals with asymptomatic infections are on average 75% as infectious as those with symptomatic infections.[9,15,16] Combined, these baseline assumptions imply that persons with infection who never develop symptoms may account for approximately 24% of all transmission (T):

$$T_{ns} = p_{ns} \times r_{ns} / (p_{ns} \times r_{ns} + [1 - p_{ns}]).$$

Table. Key Assumptions and Evidence Informing Those Assumptions

Source	Evidence base	Estimate or assumption
Assumptions for presymptomatic transmission		
Peak infectiousness relative to onset, d		
Casey et al, 2020[18]	Range, 17 studies	−3 to 1.2 d
Assumed baseline	NA	0 d
Assumed range	NA	−2 to 2 d
Assumptions for never symptomatic transmission		
Proportion never symptomatic		
Oran et al, 2020[12]	Inferred range, 16 studies	30% to 45%
Buitrago-Garcia et al, 2020[14]	Meta-estimate, 7 studies	26% to 37%
Davies et al, 2020[20]	Age-dependent estimate, 6 studies	20% to 70%
Assumed baseline	NA	30%
Relative infectiousness of individuals who never have symptoms		
Lee et al, 2020[9]	303 patients, assessment of viral shedding	Approximately 100%
Chaw et al, 2020[15]	1701 secondary contacts	40% to 140%
Mc Evoy et al, 2020[16]	Inferred range, 6 studies	40% to 70%
Assumed baseline		75%
Overall proportion of individuals who never have symptoms transmission		
Assumed range	NA	0% to 70%

Abbreviation: NA, not applicable.

We varied this overall proportion, T_{ns}, between 0% and 70% to assess a wide range of possible proportions. The daily proportion of transmission from individuals after symptom onset (T_s) was therefore:

$$T_s(t) = (1 - T_{ns}) \times p_{so}(t) \times I(t),$$

and the daily proportion of transmission from presymptomatic (T_{ps}) individuals, ie, those who develop symptoms but become infectious prior to symptom onset, is:

$$T_{ps}(t) = 1 - T_s(t) - T_{ns}.$$

We modified baseline assumptions to consider the relative importance of different levels of never symptomatic and presymptomatic transmission. Code is available in the eAppendix in the Supplement.

Statistical Analysis

All analyses were conducted in R version 4.0.1 (R Project for Statistical Computing). No statistical testing was conducted, so no prespecified level of significance was set.

Results

Under baseline assumptions, approximately 59% of all transmission came from asymptomatic transmission: 35% from presymptomatic individuals and 24% from individuals who are never symptomatic (**Figure 1**). Because each component is uncertain, we assessed different timings of peak infectiousness relative to illness onset and different proportions of transmission from individuals who never have symptoms. Maintaining the 24% of transmission from individuals who never have symptoms, but shifting peak infectiousness 1 day earlier (to day 4) increased presymptomatic transmission to 43% and all asymptomatic transmission to 67% (Figure 1A). A later peak (ie, day 6) decreased presymptomatic to 27% and all asymptomatic transmission to 51% (Figure 1C).

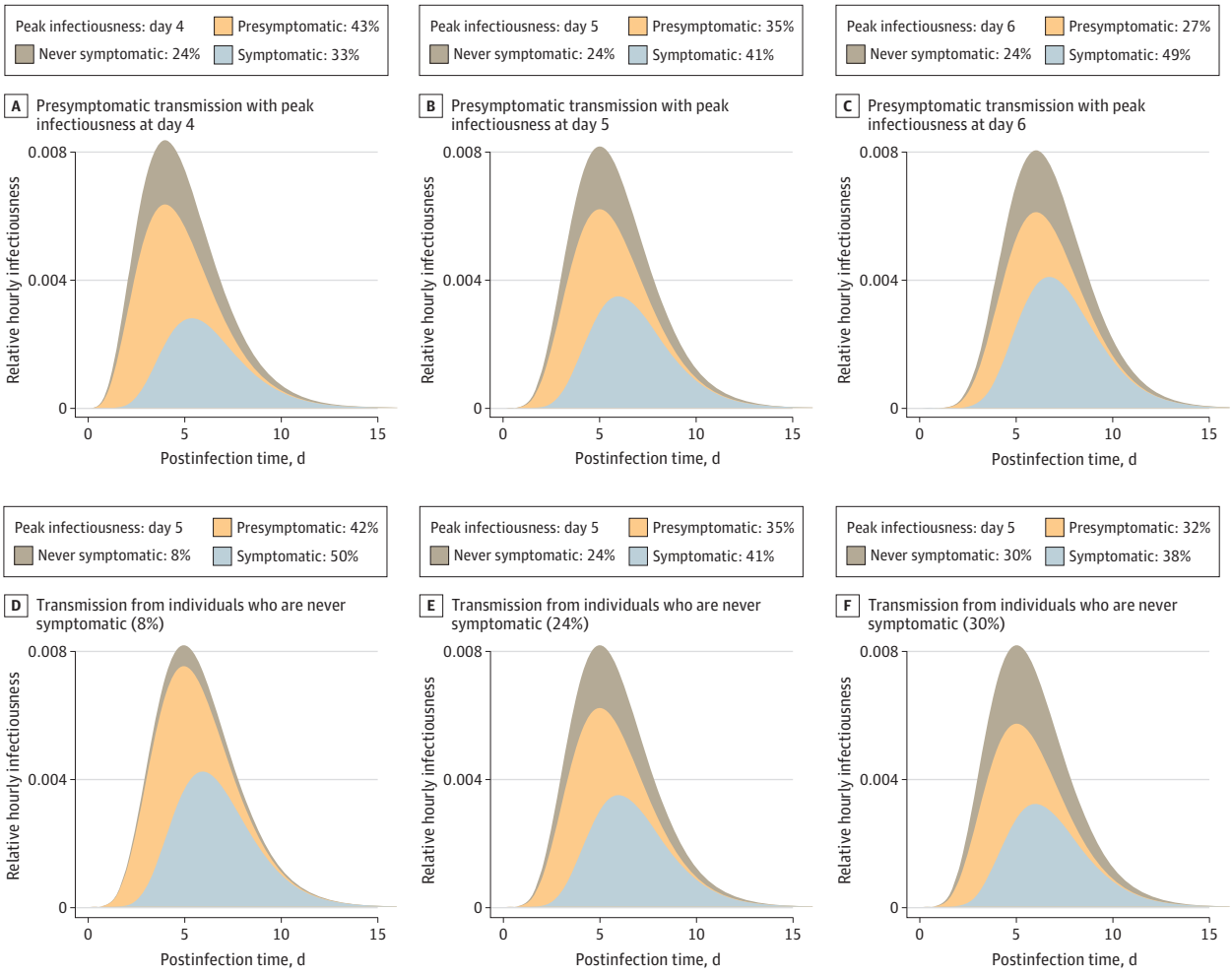
Holding the day of peak infectiousness constant at day 5 and decreasing the proportion of transmission from individuals who are never symptomatic to 10% with a relative infectiousness of 75% (baseline assumption), the proportion of all transmission from those who are never symptomatic decreased to 8%, presymptomatic transmission increased to 42%, and combined asymptomatic transmission was 50% of all transmission (Figure 1D). In contrast, if the proportion of those who ever develop symptoms was 30% and their relative infectiousness increased to 100%, they contributed 30% of all transmission, presymptomatic transmission was 32%, and combined asymptomatic transmission was 62% of all transmission (Figure 1F).

Uncertainty remains regarding the magnitude of both presymptomatic and never symptomatic transmission. Therefore, we analyzed a wider range of each of these components, with peak infectiousness varying between 2 days before (more presymptomatic transmission) to 2 days after (less presymptomatic transmission) median symptom onset and with never symptomatic transmission ranging from 0% to 70% (**Figure 2**). Under this broader range of scenarios, most combined assumptions of peak infectiousness timing and transmission from individuals who never have symptoms indicated that at least 50% of new SARS-CoV-2 infections likely originated from individuals without symptoms at the time of transmission. If more than 30% of transmission was from individuals who never have symptoms, total asymptomatic transmission was higher than 50% with any value of peak infectiousness, up to 2 days after the median time of symptom onset. If peak infectiousness was at any point approximately 6 hours before median symptom onset time, more than 50% of transmission was from individuals without symptoms, regardless of the proportion from those who never have symptoms. Even a very conservative assumption of peak infectiousness 2 days post–median onset and 0% never symptomatic transmission still resulted in more than 25% of transmission from asymptomatic individuals.

Discussion

The findings presented here complement an earlier assessment[21] and reinforce the importance of asymptomatic transmission: across a range of plausible scenarios, at least 50% of transmission was estimated to have occurred from persons without symptoms. This overall proportion of transmission from presymptomatic and never symptomatic individuals is key to identifying mitigation measures that may be able to control SARS-CoV-2. For example, if the reproduction number (*R*) in a given setting is 2.0, then at least a 50% reduction in transmission is needed to drive the reproductive number below 1.0. Given that in some settings *R* is likely much greater than 2 and more than half of transmissions may come from individuals who are asymptomatic at the time of transmission, effective control must mitigate transmission risk from people without symptoms.

Figure 1. The Contribution of Asymptomatic Transmission Under Different Infection Profiles



The top curve in each panel represents the average relative hourly infectiousness, such that while the lower curves change under different assumptions, the total hourly infectiousness equals 1 in all cases. Within each curve, the colored area indicates the proportion of transmission from each class of individuals. The portion attributed to individuals with symptoms (light blue) can also be interpreted as the maximum proportion of transmission that can be controlled by immediate isolation of all symptomatic cases. Panels A, B, and C show different levels of presymptomatic transmission. We calibrated infectiousness to peak at day 4 (A), 5 (B; median incubation period), or 6 (C) days. Panels D, E, and F show different proportions of transmission from individuals who are never symptomatic: 8% (C; eg, 10% never symptomatic and 75% relative infectivity), 24% (D; baseline, 30% never symptomatic and 75% relative infectivity), and 30% (E; eg, 30% never symptomatic and 100% relative infectivity).

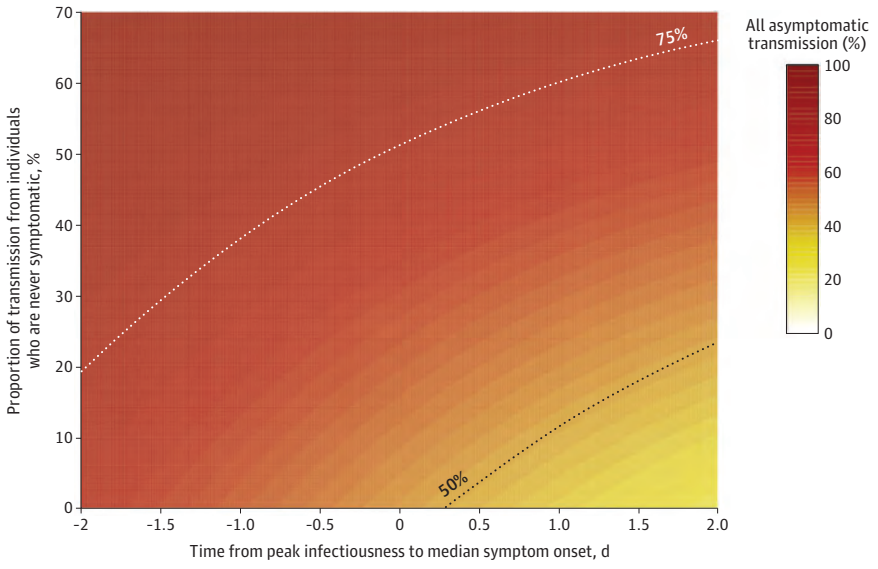
Limitations

This study has limitations. First, we used a simplistic model to represent a complex phenomenon, ie, the average infectiousness of SARS-CoV-2 infections over time. We used this model deliberately to test assumptions about the timing of peak infectiousness and transmission among asymptomatic individuals so that we could vary only these 2 critical parameters and assess their relative effects. Therefore, these results lack quantitative precision, but they demonstrate the qualitative roles of these 2 components and show that across broad ranges of possible assumptions, the finding that asymptomatic transmission is a critical component of SARS-CoV-2 transmission dynamics remains constant.

As discussed here, the exact proportions of presymptomatic and never symptomatic transmission are not known. This also applies to the incubation period estimates, which are based on individual exposure and onset windows that are difficult to observe with precision and therefore include substantial uncertainty even when leveraging estimates across multiple studies. Moreover, they likely vary substantially in different populations. For example, older individuals are more likely than younger persons to experience symptoms,[20] so in populations of older individuals, never asymptomatic transmission may be less important. However, specific age groups are rarely exclusively isolated from other age groups, so asymptomatic transmission risk is still important in those groups and even more so in younger age groups, in which transmission may be even more dominated by asymptomatic transmission.[20]

Real-world transmission dynamics are also not entirely dependent on the individual-level dynamics of infectiousness over time. Now that COVID-19 is widely recognized, individuals with COVID-19 symptoms are more likely to isolate themselves and further reduce the proportion of transmission from symptomatic individuals, shifting a greater proportion of transmission to those who do not have symptoms. In this sense, the estimates here represent the lower end of the proportion of asymptomatic transmission in the presence of interventions to reduce symptomatic transmission.

Figure 2. Combined Transmission From Individuals Who Are Presymptomatic and Those Who Never Have Symptoms



Colors indicate the proportion of transmission due to all individuals without symptoms at the time of transmission, including presymptomatic transmission (x-axis, the timing of peak infectiousness relative to symptom onset) and transmission from individuals who are never symptomatic (y-axis). For example, peak infectiousness at the same time as median symptom onset (0 days difference) with 10% of transmission from individuals who never have symptoms would mean that approximately 51% of transmission is from asymptomatic individuals.

Conclusions

Under a range of assumptions of presymptomatic transmission and transmission from individuals with infection who never develop symptoms, the model presented here estimated that more than half of transmission comes from asymptomatic individuals. In the absence of effective and widespread use of therapeutics or vaccines that can shorten or eliminate infectivity, successful control of SARS-CoV-2 cannot rely solely on identifying and isolating symptomatic cases; even if implemented effectively, this strategy would be insufficient. These findings suggest that effective control also requires reducing the risk of transmission from people with infection who do not have symptoms. Measures such as mask wearing and social distancing empower individuals to protect themselves and, if infected, to reduce risk to their communities.[21] These measures can also be supplemented by strategic testing of people who are not ill, such as those who have exposures to known cases (eg, contact tracing) or are at high risk of exposing others (eg, congregate facility staff, those with frequent contact with the public). Multiple measures that effectively address transmission risk in the absence of symptoms are imperative to control SARS-CoV-2.

ARTICLE INFORMATION

Accepted for Publication: December 7, 2020.

Published: January 7, 2021. doi:10.1001/jamanetworkopen.2020.35057

Correction: This article was corrected on February 12, 2021, to fix an error in the Supplement.

Corresponding Author: Jay C. Butler, MD, Office of the Deputy Director for Infectious Diseases, US Centers for Disease Control and Prevention, 1600 Clifton Rd, Mailstop H24-12, Atlanta, GA 30329 (jcb3@cdc.gov).

Author Affiliations: COVID-19 Response, US Centers for Disease Control and Prevention, Atlanta, Georgia (Johansson, Quandelacy, Kada, Prasad, Steele, Brooks, Slayton, Biggerstaff); Office of the Deputy Directory for Infectious Diseases, US Centers for Disease Control and Prevention, Atlanta, Georgia (Johansson, Slayton, Biggerstaff, Butler).

Author Contributions: Dr Johansson had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Johansson, Quandelacy, Kada, Brooks, Slayton, Butler.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Johansson, Quandelacy, Brooks, Biggerstaff, Butler.

Critical revision of the manuscript for important intellectual content: Johansson, Kada, Prasad, Steele, Brooks, Slayton, Biggerstaff, Butler.

Statistical analysis: Johansson, Quandelacy, Kada.

Administrative, technical, or material support: Prasad, Steele, Brooks, Biggerstaff, Butler.

Supervision: Johansson, Butler.

Conflict of Interest Disclosures: None reported.

Funding/Support: This work was performed as part of the US Centers for Disease Control and Prevention's coronavirus disease 2019 response and was supported solely by federal base and response funding.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

REFERENCES

1. Peiris JS, Yuen KY, Osterhaus AD, Stöhr K. The severe acute respiratory syndrome. *N Engl J Med*. 2003;349(25):2431-2441. doi:10.1056/NEJMra032498

2. Tindale LC, Stockdale JE, Coombe M, et al. Evidence for transmission of COVID-19 prior to symptom onset. *Elife*. 2020;9:e57149. doi:10.7554/eLife.57149

3. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. *Int J Infect Dis*. 2020;93:284-286. doi:10.1016/j.ijid.2020.02.060

4. Zhao S, Gao D, Zhuang Z, et al. Estimating the serial interval of the novel coronavirus disease (COVID-19): a statistical analysis using the public data in Hong Kong from January 16 to February 15, 2020. *Front Phys*. Published online September 17, 2020. doi:10.3389/fphy.2020.00347

5. Du Z, Xu X, Wu Y, Wang L, Cowling BJ, Meyers LA. Serial interval of COVID-19 among publicly reported confirmed cases. *Emerg Infect Dis*. 2020;26(6):1341-1343. doi:10.3201/eid2606.200357

6. Wei WE, Li Z, Chiew CJ, Yong SE, Toh MP, Lee VJ. Presymptomatic transmission of SARS-CoV-2—Singapore, January 23-March 16, 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(14):411-415. doi:10.15585/mmwr.mm6914e1

7. Tong Z-D, Tang A, Li K-F, et al. Potential presymptomatic transmission of SARS-CoV-2, Zhejiang Province, China, 2020. *Emerg Infect Dis*. 2020;26(5):1052-1054. doi:10.3201/eid2605.200198

8. He X, Lau EHY, Wu P, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med*. 2020;26(5):672-675. doi:10.1038/s41591-020-0869-5

9. Lee S, Kim T, Lee E, et al. Clinical course and molecular viral shedding among asymptomatic and symptomatic patients with SARS-CoV-2 infection in a community treatment center in the Republic of Korea. *JAMA Intern Med*. 2020. doi:10.1001/jamainternmed.2020.3862

10. Benefield AE, Skrip LA, Clement A, Althouse RA, Chang S, Althouse BM. SARS-CoV-2 viral load peaks prior to symptom onset: a systematic review and individual-pooled analysis of coronavirus viral load from 66 studies. *medRxiv*. Preprinted published online September 30, 2020. doi:10.1101/2020.09.28.20202028

11. Byrne AW, McEvoy D, Collins AB, et al. Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases. *BMJ Open*. 2020; 10(8):e039856. doi:10.1136/bmjopen-2020-039856

12. Oran DP, Topol EJ. Prevalence of asymptomatic SARS-CoV-2 infection : a narrative review. *Ann Intern Med*. 2020;173(5):362-367. doi:10.7326/M20-3012

13. Poletti P, Tirani M, Cereda D, et al. Probability of symptoms and critical disease after SARS-CoV-2 infection. arXiv. Preprint published online June 22, 2020. Accessed December 10, 2020. https://arxiv.org/abs/2006.08471

14. Buitrago-Garcia D, Egli-Gany D, Counotte MJ, et al. Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: a living systematic review and meta-analysis. *PLoS Med*. 2020;17(9): e1003346. doi:10.1371/journal.pmed.1003346

15. Chaw L, Koh WC, Jamaludin SA, Naing L, Alikhan MF, Wong J. Analysis of SARS-CoV-2 transmission in different settings, Brunei. *Emerg Infect Dis*. 2020;26(11):2598-2606. doi:10.3201/eid2611.202263

16. Mc Evoy D, McAloon CG, Collins AB, et al. The relative infectiousness of asymptomatic SARS-CoV-2 infected persons compared with symptomatic individuals: a rapid scoping review. *medRxiv*. Preprint published online August 1, 2020. doi:10.1101/2020.07.30.20165084

17. McAloon C, Collins Á, Hunt K, et al. Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research. *BMJ Open*. 2020;10(8):e039652. doi:10.1136/bmjopen-2020-039652

18. Casey M, Griffin J, McAloon CG, et al. Pre-symptomatic transmission of SARS-CoV-2 infection: a secondary analysis using published data. *medRxiv*. Preprint published online June 11, 2020. doi:10.1101/2020.05.08.20094870

19. Byambasuren O, Dobler CC, Bell K, et al. Comparison of seroprevalence of SARS-CoV-2 infections with cumulative and imputed COVID-19 cases: systematic review. *medRxiv*. Preprint published online October 22, 2020. doi:10.1101/2020.07.13.20153163

20. Davies NG, Klepac P, Liu Y, Prem K, Jit M, Eggo RM; CMMID COVID-19 working group. Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat Med*. 2020;26(8):1205-1211. doi:10.1038/s41591-020-0962-9

21. US Centers for Disease Control and Prevention. Things to know about the COVID-19 pandemic. Updated December 4, 2020. Accessed December 10, 2020. https://www.cdc.gov/coronavirus/2019-ncov/your-health/need-to-know.html

SUPPLEMENT.

eAppendix. Code for Analysis

Exhibit

42

SHOPP
Society for Healthcare Organization Procurement Professionals

April 7th, 2020

To all post-acute care provider advocates:

SHOPP (The Society for Healthcare Organization Procurement Professionals) is an independent, non-profit entity created to improve quality and efficiency in post-acute care. We do not sell or lobby, we advocate for the provider and seek nothing in return.

To assist in your efforts to secure relief funding, we compiled this analysis of emergency, unfunded, marginal PPE costs incurred by skilled nursing facilities and assisted living centers treating COVID-19 patients. This is not an estimate, our calculations are based on current market pricing and CDC guidelines, per CMS' April 2nd, 2020 guidance.

For each item, SHOPP analyzed the pre–COVID-19 cost as well as the current COVID-19 pricing as of April 6th, 2020. (It is important to note that COVID-19 pricing has been extremely volatile and continues to change daily.) SHOPP has supplied both the price markup and percentage markup per item. Each item was reviewed with Faygee Morgenshtern, CEO of People Powered Nursing and Michael Greenfield, CEO of Prime Source Healthcare Solutions and Co-Founder of SHOPP. This analysis was conducted using a 100-bed facility multiplied by 30 days for a total of 3,000 census days per month. The items evaluated were:

- Vinyl exam gloves
- Latex gloves
- Nitryl gloves
- 3-ply masks
- KN95 masks
- N95 masks
- 3M N95 masks
- Hand sanitizer
- Isolation gowns
- Face shields
- Soap

When calculating the PPD (patient per day) cost, SHOPP took into consideration that a given facility does not use all the evaluated items; instead, facilities generally use one type of glove and one type of mask. As an example, facilities ideally use vinyl exam gloves daily. If vinyl is not available, the next choice from a price perspective is latex gloves, but many people are allergic to latex. Nitryl gloves are the most expensive option, but in circumstances like COVID-19, they are being used more frequently.

Society for Healthcare Organization Procurement Professionals

3-ply mask pricing was calculated using FDA-approved masks in boxes rather than bags, and N95 mask pricing was calculated using NIOSH-certified masks on the CDC-approved list. It is important to note that KN95 masks were only FDA-approved as of April 2nd, 2020. For 3-ply masks, SHOPP's calculation assumed every employee in the facility (both medical and non-medical staff) would be wearing 2 masks a day. KN95, N95, and 3M N95 masks, as well as face shields, last longer and can be used for a 7-day period. (At the time of our analysis, 3M N95 masks were currently unavailable) The calculation for these masks and for face shields assumed supplies for 175 employees to account for both PRN staff and regular staff. Isolation gowns were calculated as being changed after every resident encounter similar to gloves. It is important to note that assisted living centers use roughly 25% less PPE than skilled nursing facilities based on resident interaction.

For gloves, regardless of type, SHOPP calculated that all medical and non-medical staff change gloves between each resident as well as twice a day for non-resident care (bathroom or food breaks, as well as any personal breaks). Lastly, for soap and hand sanitizer, the calculation assumed each employee would use 2 ounces of each daily. Hand sanitizer pricing was calculated on a per-ounce basis for an 8-ounce bottle. Soap pricing was calculated on a per-ounce basis via a dispenser.

Calculating the PPD cost per item, SHOPP first reviewed pre–COVID-19 pricing while following pre–COVID-19 requirements for PPE (i.e., vinyl gloves and soap). The PPD costs for this equipment totaled $35.00 per day or 0.35 PPD. The next breakdown was calculated using pre–COVID-19 pricing while following current CDC guidelines for PPE (i.e. vinyl gloves N95 masks, hand sanitizer, isolation gowns, face shields and soap). The total cost per day was $236.00 or 2.36 PPD. This is a 674% increase in costs.

Finally, SHOPP's analysis examined current COVID-19 pricing following the current CDC guidelines. This was broken out into two categories: typical PPD using vinyl gloves and N95 masks, and PPD using nitryl gloves (due to unavailability of vinyl gloves) and N95 masks. For typical PPD, the daily cost was $2,510.25 or 25.10 PPD. This is a 1,064% increase in costs compared to the pre–COVID-19 costs and requirements. For PPD using nitryl gloves, the daily cost was $2,558.25 or 25.58 PPD. This is a 1,084% increase in costs.

End of report narrative.

Thanks,

CO - Founders

Marc Zimmet Josh Silverberg Ari Stawis Michael Greenfield

Supplemental Notes

Exhibit 1: Pricing Per Item*

Item	Pre Covid-19 Cost	Current Covid-19 Cost	Price Markup	Percentage Markup	Notes
Vinyl Exam Gloves	$ 0.02	$ 0.06	$ 0.04	300%	Ideally people use vinyl. If not vinal then latex but issues with allergies so Nitryl is most expensive resort
Latex Gloves	$ 0.03	$ 0.08	$ 0.05	267%	
Nitryl Gloves	$ 0.05	$ 0.10	$ 0.05	200%	
3ply Masks	$ 0.05	$ 0.75	$ 0.70	1500%	FDA certified in boxes, not bags.
KN95 Masks	N/A	$ 4.00		N/A	Recently (April 2nd) FDA approved
N95 Masks	$ 0.38	$ 5.75	$ 5.37	1513%	Niosh certified, CDC approved list
3M N95 Masks	$ 0.11	$ 6.75	$ 6.64	6136%	Currently unavailable. Includes testing kit for sizing
Hand Sanitizer	$ 0.26	$ 0.56	$ 0.30	215%	Per ounce for 8 ounce bottle. Required to be at 70% percent alcohol
Isolation Gowns	$ 0.25	$ 5.00	$ 4.75	2000%	New gown per resident
Face shields	$ 0.50	$ 4.50	$ 4.00	900%	Reusable
Soap	$ 0.19	$ 0.35	$ 0.16	184%	Per ounce via dispenser

Exhibit 2: PPD Formulas*

Item	Old	Old Pricing w/ New Demand	CDC Guidelines	Notes
Vinyl Exam Gloves	$ 16.00	$ 24.00	$ 72.00	All staff changing between each resident as well as 2x/day for non-resident care
Latex Gloves	$ 96.00	$ 96.00	$ 96.00	All staff changing between each resident as well as 2x/day for non-resident care
Nitryl Gloves	$ 60.00	$ 60.00	$ 120.00	All staff changing between each resident as well as 2x/day for non-resident care
3-Ply Masks	$ 10.00	$ 10.00	$ 150.00	100 employees x 2 masks a day
KN95 Masks	$ 100.00	$ 100.00	$ 100.00	175 employees x 4 dollars a mask divided by 7 days
N95 Masks	$ 9.50	$ 9.50	$ 143.75	175 employees x 5.75 dollars a mask divided by 7 days
3M N95 Masks	$ 168.75	$ 168.75	$ 168.75	175 employees x 6.75 dollars a mask divided by 7 days
Hand Sanitizer	$ 52.00	$ 52.00	$ 112.00	Each employee using 2 ounce a day
Isolation Gowns	$ 100.00	$ 100.00	$ 2,000.00	100 residents x 4 visits a day x price per gown
Face Shields	$ 12.50	$ 12.50	$ 112.50	175 employees x 4.50 per face shield divided by 7 days
Soap	$ 19.00	$ 38.00	$ 70.00	100 employees x 2 ounce per Employee

Exhibit 3: PPD Costs*

	Pre-COVID-19 Requirements	Pre-COVID-19 Pricing w/ Current CDC Guidelines	Current COVID-19 Pricing w/ Current CDC Guidelines	Current COVID-19 Pricing w/ Current CDC Guidelines, Nitryl Gloves, N95 Masks
Total Costs Per Day	$ 35.00	$ 236.00	$ 2,510.25	$ 2,558.25
PPD	$ 0.35	$ 2.36	$ 25.10	$ 25.58
Percentage Markup		674%	1064%	1084%

*Pre-Covid-19 pricing is based on multiple facilities averaged over 12-month period.

Exhibit

43

Article

Surveillance of SARS-CoV-2 at the Huanan Seafood Market

https://doi.org/10.1038/s41586-023-06043-2

Received: 17 February 2022

Accepted: 3 April 2023

Published online: 5 April 2023

Check for updates

William J. Liu[1,7]✉, Peipei Liu[1,7], Wenwen Lei[1,7], Zhiyuan Jia[1,7], Xiaozhou He[1,7], Weifeng Shi[2,7], Yun Tan[3,7], Shumei Zou[1], Gary Wong[4], Ji Wang[1], Feng Wang[1], Gang Wang[1], Kun Qin[1], Rongbao Gao[1], Jie Zhang[1], Min Li[1], Wenling Xiao[1,5], Yuanyuan Guo[1], Ziqian Xu[1], Yingze Zhao[1], Jingdong Song[1], Jing Zhang[1], Wei Zhen[1], Wenting Zhou[1], Beiwei Ye[1], Juan Song[1], Mengjie Yang[1], Weimin Zhou[1], Yuting Dai[3], Gang Lu[3], Yuhai Bi[6], Wenjie Tan[1], Jun Han[1], George F. Gao[1,6]✉ & Guizhen Wu[1]✉

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of coronavirus disease 2019, emerged in December 2019. Its origins remain uncertain. It has been reported that a number of the early human cases of coronavirus disease 2019 had a history of contact with the Huanan Seafood Market. Here we present the results of surveillance for SARS-CoV-2 within the market. From 1 January 2020, after closure of the market, 923 samples were collected from the environment. From 18 January, 457 samples were collected from 18 species of animal, comprising unsold contents of refrigerators and freezers, swabs from stray animals and the contents of a fish tank. Using quantitative real-time polymerase chain reaction (RT-qPCR) and high-throughput sequencing (Bowtie2 analysis), SARS-CoV-2 was detected in 74 (70 RT-qPCR and 4 Bowtie2) environmental samples, but none of the animal samples. Three live viruses were successfully isolated. The viruses from the market shared a nucleotide identity of 99.99% to 100% with the human isolate HCoV-19/Wuhan/IVDC-HB-01/2019. SARS-CoV-2 lineage A (8782T and 28144C) was found in an environmental sample. RNA-sequencing analysis of SARS-CoV-2-positive and SARS-CoV-2-negative environmental samples showed an abundance of different vertebrate genera at the market. In summary, this study provides information about the distribution and prevalence of SARS-CoV-2 in the Huanan Seafood Market during the early stages of the outbreak of coronavirus disease 2019.

Infections with human coronavirus 2019 (HCoV-19)[1,2], named as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the International Committee on Taxonomy of Viruses[3], can result in coronavirus disease 2019 (COVID-19), characterized by various clinical outcomes from asymptomatic infections to severe pneumonia and even death[4,5]. Globally, as of 28 February 2023, more than 758 million confirmed cases and more than 6.8 million deaths have been reported (https://covid19.who.int).

Human individuals with COVID-19 were first reported in late December 2019, in Wuhan, China, as pneumonia of unknown aetiology. A certain proportion of of these early cases were found to be linked to the Huanan Seafood Market (HSM) in Wuhan[4,6], where various animal meats, exotic seafood and live animals were available for purchase. The HSM has been suspected to be the source of the COVID-19 pandemic[7]. Not all of the early human cases had epidemiological links to the market[6,8], and alternative hypotheses for the market association (for example, entry of the virus into the market through humans or the cold chain) also exist.

SARS-CoV-2 has high similarity to a few coronaviruses derived from bats in Asian countries including China, Laos, Japan, Cambodia and Thailand, and some scientists have proposed that bats might be the original source of SARS-CoV-2 (refs. 1,8–14). Whether another animal might have acted as an intermediate host to facilitate virus spillover from bats to humans is still unknown[15,16]. An important finding was the discovery of SARS-CoV-2-related coronaviruses from pangolins, in which the spike proteins contained receptor-binding domains showing high similarity to the receptor-binding domain of SARS-CoV-2 (refs. 17–19). Pangolins might be involved in the ecology of coronaviruses, but whether they are the intermediate host for SARS-CoV-2 is unknown, given the current data[20]. A recent study documented the animal species in the HSM between May 2017 and November 2019 and noted that no pangolins or bats were present, but some animals proposed to be susceptible to

[1]NHC Key Laboratory of Biosafety, National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention (China CDC), Beijing, China. [2]Key Laboratory of Emerging Infectious Diseases in Universities of Shandong, Shandong First Medical University, and Shandong Academy of Medical Sciences, Tai'an, China. [3]Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics, National Research Center for Translational Medicine, Ruijin Hospital Affiliated to Shanghai Jiao Tong University (SJTU) School of Medicine, Shanghai, China. [4]CAS Key Laboratory of Molecular Virology & Immunology, Institut Pasteur of Shanghai, Chinese Academy of Sciences (CAS), Shanghai, China. [5]School of Laboratory Medicine and Life Sciences, Wenzhou Medical University, Wenzhou, China. [6]CAS Key Laboratory of Pathogen Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences (CAS), Beijing, China. [7]These authors contributed equally: William J. Liu, Peipei Liu, Wenwen Lei, Zhiyuan Jia, Xiaozhou He, Weifeng Shi, Yun Tan. ✉e-mail: liujun@ivdc.chinacdc.cn; gaofu@chinacdc.cn; wugz@ivdc.chinacdc.cn

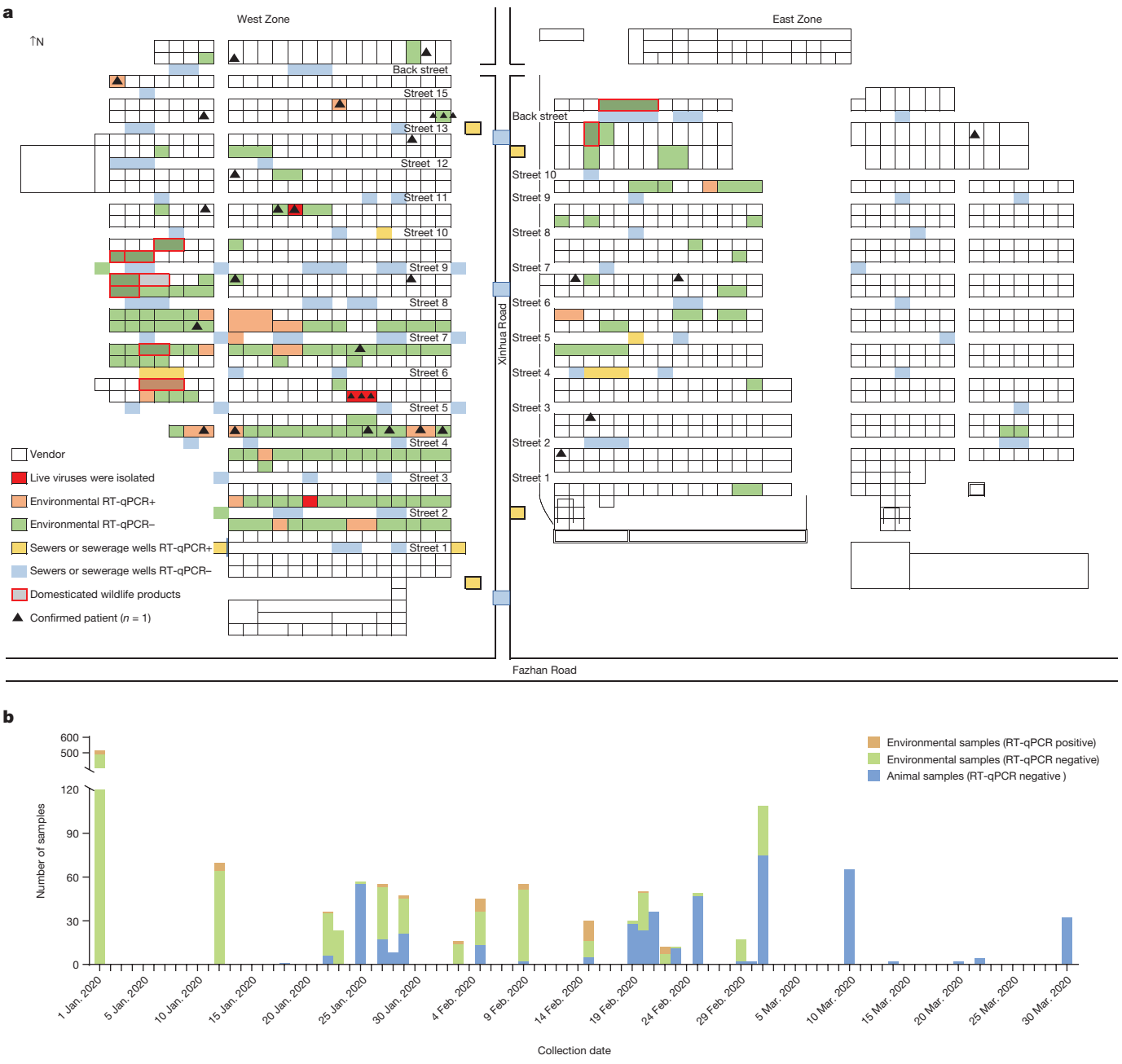


Fig. 1 | The distribution of the positive environmental samples in the HSM. **a**, As the place of the early cluster of patients with COVID-19, the HSM is separated into East and West Zones with the Xinhua Road between them. To detect the presence of SARS-CoV-2 RNA, RT-qPCR was carried out. The locations of the positive samples are marked in the map of the market in orange, and the locations of the samples that the live viruses were isolated from are filled red. The map also shows locations of stalls where domesticated wildlife products were sold. **b**, Timeline of environmental and animal samples collected within and around the HSM. The data for confirmed cases up to 31 December 2019 were taken from ref. 24.

sarbecoviruses, such as raccoon dogs, were present[21]. Thus far, the origins of SARS-CoV-2 (refs. 22,23) and the role of the HSM in the origins and spread of SARS-CoV-2 remain unclear. The data from the HSM may provide important information.

The HSM is located in the Jianghan District in the downtown area of Wuhan, the capital city of Hubei Province, and is approximately 800 m away from Hankou Railway Station, a major railway travel hub. It occupies >50,000 m^2, with 678 stalls located close to each other in extremely crowded conditions (Fig. 1a). The market is separated into two zones, the East and West Zones, with seafood and animals mainly sold in the West Zone and livestock meat sold in the East Zone. Among the 678 stalls of the market, 10 stalls selling domesticated wildlife (1.5%) were identified according to sale records[24], located in the southwestern corner of the West Zone (8/10) and the northwestern corner of the East Zone (2/10; Fig. 1a). According to sale records, during late December 2019, animals or animal products were sold in these 10 animal stalls. Animals included snakes, avian species (chickens, ducks, geese, pheasants and doves), sika deer, badgers, rabbits, bamboo rats, porcupines, hedgehogs, salamanders, giant salamanders, bay crocodiles, Siamese crocodiles and so on, among which snakes, salamanders and crocodiles were traded as live animals (described in detail in ref. 24).

The market was closed in the morning of 1 January 2020, shortly after the identification of the pneumonia of unknown aetiology. On the same day, in the early morning, the Chinese Center for Disease

Control and Prevention (China CDC) dispatched an epidemiological team, together with experts from Hubei Provincial CDC and Wuhan Municipal CDC, to the HSM to collect environmental samples and study the potential introduction of SARS-CoV-2 into the market (Fig. 1b). From 1 January 2020 until 2 March 2020, a total of 923 environmental samples from different locations within and around the market and 457 animal samples, including dead animals in refrigerators and freezers and stray animals and their faeces, were collected, with some stray animals sampled until 30 March (Extended Data Tables 1–3 and Supplementary Table 1). After the closure of the market, the outside surface of the rolling shutter doors of the stalls and the corridors were disinfected (with 1% bleach mixed with water) throughout January and February 2020. The goods inside the stalls were completely cleared and disinfected until early March 2020.

Out of the 923 environmental samples collected in and around the HSM, 74 were found by the quantitative real-time polymerase chain reactions (RT-qPCR, 70 positive samples) and high-throughput sequencing (Bowtie2 analysis, 4 positive samples with non-3′ poly-A reads) to be positive for SARS-CoV-2 with a positivity rate of 8.0%. Cycle threshold (Ct) values for the RT-qPCR ranged from 23.9 to 41.7 (Supplementary Table 2). Among the 828 samples from inside the HSM, 64 samples (7.7%) were positive. Of the 64 SARS-CoV-2-positive samples collected inside the HSM, 87.5% (56/64) were collected in the West Zone of the market, particularly in streets 1 to 8, with 71.4% (40/56) positive samples identified herein (Fig. 1a). Among the 14 samples from warehouses related to the HSM, 5 tested positive. This may reflect the nature of SARS-CoV-2 presence in the market during the early phase of the outbreak. Among the 51 samples from sewerage wells (Supplementary Table 1) in the surrounding areas outside the HSM, 4 tested positive (Supplementary Table 2). Notably, 1 sample (Env_0601), a floor surface swab, out of the 30 environmental samples collected from Dongxihu Market in Wuhan on 22 January 2020 also tested positive (Supplementary Table 2 and Extended Data Table 4).

Of the 110 samples collected from sewers or sewerage wells in the market, 24 samples were positive for SARS-CoV-2 nucleic acid. All four sewerage wells in the market tested positive. During the onsite investigation of the overground drainage pathway in the HSM, we found that the wastewater in the overground drainage led into the underground drainage inside the market and then flowed into the wells on the edge of the market. We then did a spot-check sampling across all of the overground drains according to the principles described in the Methods (Extended Data Fig. 1). Excreta of the upper respiratory tract of infected humans and the potential animal waste would be mixed together into the overground drainage. Thus, these data suggested either that infected people and/or animals in the market contaminated the sewage or that the contaminated sewage may have had a role in furthering the virus transmission within the case cluster in the market.

The merchants' activities were assessed against the RT-qPCR results of the environmental samples. The sampling covered 19.8% (134/678) of the shops in the market (95% confidence interval (CI): 16.8–23.0%). Of the positive samples, 44 were distributed among 21 shops in the market, 19 of whom were located in the West Zone with the remaining 2 located in the East Zone (Fig. 1a). Some vendors sold more than one type of product. Although the results provided some indication of an association of cases with different products, no significant differences were observed between different types of shop, including those selling poultry (22%, 8/37, 95% CI: 9.8–38.2%), cold-chain products (18.4%, 16/87, 95% CI: 10.9–28.1%), aquatic products (17.8%, 13/73, 95% CI: 9.8–28.5%), livestock (14%, 5/36, 95% CI: 4.7–29.5%), seafood products (11%, 6/56, 95% CI: 4–21.9%), wildlife products (11%, 1/9, 95% CI: 0.3–48.2%) and vegetables (25%, 2/8, 95% CI: 3.2–65%; Extended Data Fig. 2 and Extended Data Table 5). The detection of SARS-CoV-2 in several shops selling different product types suggested that SARS-CoV-2 may have been circulating in the market, especially in the West Zone, for a while in December 2019, leading to an extensive distribution of the virus within the market, which may have been facilitated by the crowded buyers and the contaminated environment.

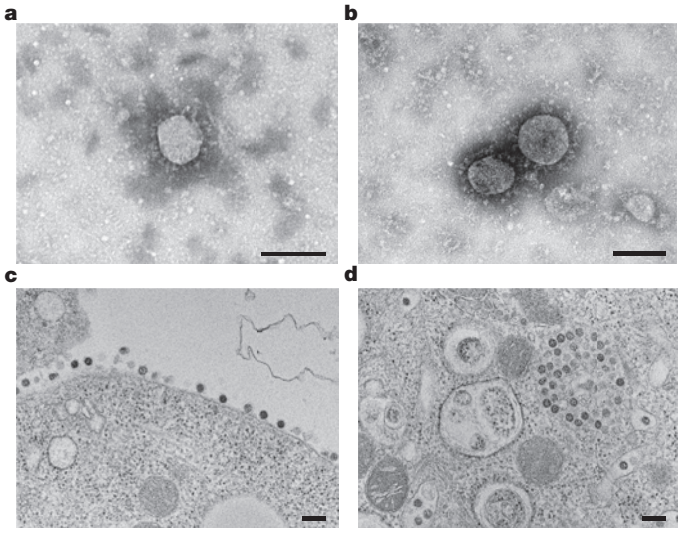


Fig. 2 | SARS-CoV-2 virus isolation from environmental samples of the HSM. **a**–**d**, Electron micrographs of the SARS-CoV-2 viruses isolated from the environmental samples in the HSM. To determine whether SARS-CoV-2 particles could be visualized from the cell supernatant and lysate, we used transmission electron microscopy to observe the culture supernatant and ultrathin sections from Vero E6 and Huh7.5 cells. The electron micrographs showed that virus particles were present in both the supernatant (**a**,**b**) and the cells (**c**,**d**). Negative-stained virus particles were generally spherical, pleomorphic and 60–140 nm in diameter. Spike protrusions were observed around the particles in a crown (corona) shape (**a**,**b**). In ultrathin cultured cell sections, a group of virus particles can be seen outside the cell (**c**), and sheets of virus particles can also be observed inside the cells (**d**). The micrographs are representatives of repeated experiments.

The 457 animal samples included 188 individuals belonging to 18 species (with some stray animals sampled until 30 March; Extended Data Table 6). The sources of the samples included unsold goods kept in refrigerators and freezers in the stalls of the HSM, and goods kept in warehouses and refrigerators related to the HSM. Three Chinese giant salamanders, which were found in a fish tank, were alive and swab samples were collected and tested. Samples from stray animals in the market were also collected, comprising swab samples from 10 cats, 27 samples of cat faeces, 1 dog, 1 weasel and 10 rats. All of the 457 animal samples tested negative for SARS-CoV-2 nucleic acid.

To determine whether there was live virus in the HSM, we inoculated 27 SARS-CoV-2-positive environmental samples collected on 1 January 2020, into cell lines, including Vero E6 and Huh7.5 cells. Cytopathic effects were observed 3 days post inoculation with sample Env_0313 on Vero E6 cells. Cytopathic effects were also observed 5 days post inoculation on Huh7.5 cells. The electron micrographs of Vero E6 cells at 5 days post inoculation showed that virus particles were present in both the supernatant and the cells. Negative-stained virus particles and ultrathin cultured cell sections showed typical coronavirus morphology (Fig. 2). Live viruses were isolated from samples Env_0313, Env_0354 and Env_0126, which were the only three samples with Ct values < 30 in the RT-qPCR. Env_0354 and Env_0126 were two swab samples from the ground and Env_0313 was swab samples from a wall. Notably, samples Env_0313 and Env_0126 were from stalls with confirmed cases. All of the results of successful virus isolation from the original samples with low Ct values revealed the existence of live SARS-CoV-2 with high titres in the environment of the HSM. Owing to the high Ct values, we did not attempt virus isolation from the samples collected at later time points.

During later sampling in the HSM in February, we collected samples to investigate the virus RNA persistence in the market. Some of these

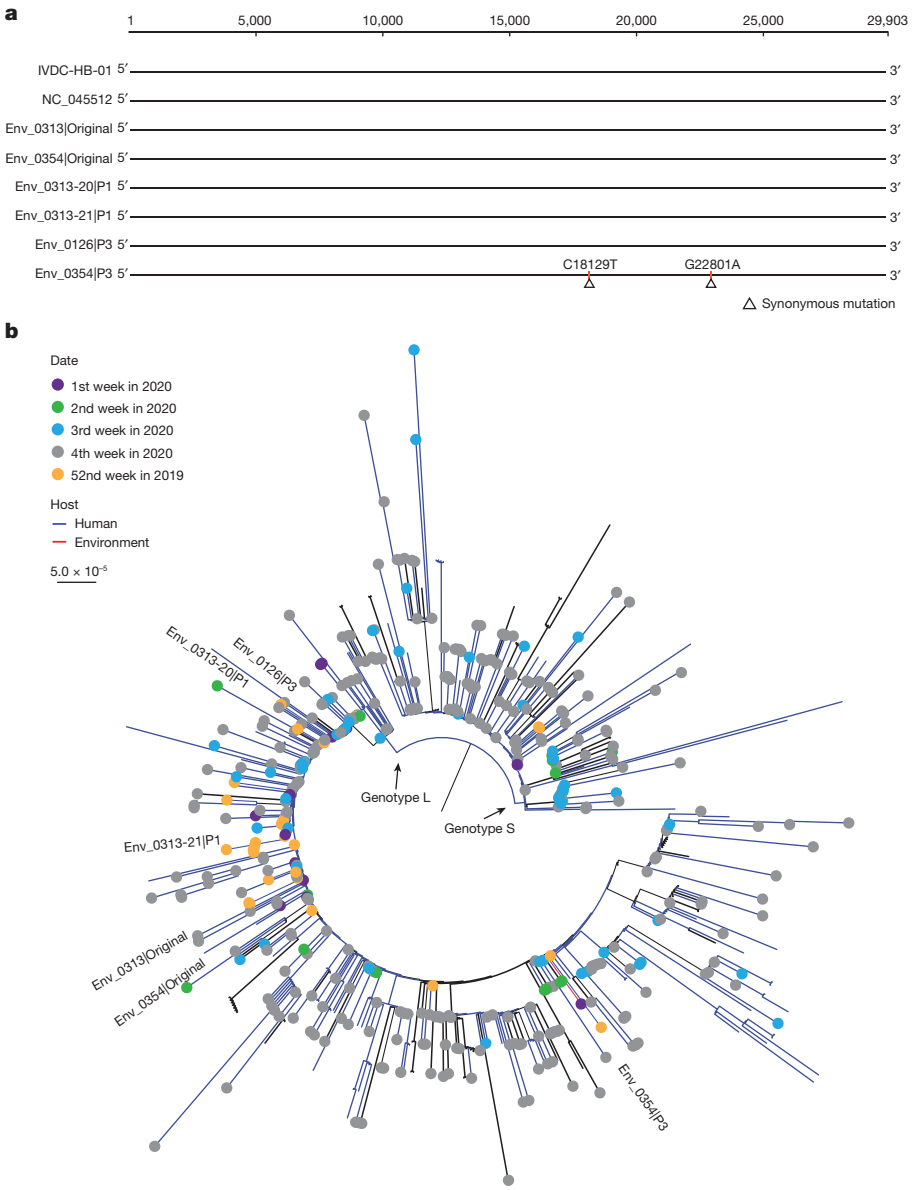


Fig. 3 | Genomic and phylogenetic analyses of SARS-CoV-2 virus genomes from the HSM. a, Sequence comparison of the full-length SARS-CoV-2 genomes in the environmental samples. **b**, Phylogenetic analysis of full-length SARS-CoV-2 genomes from the HSM and representative strains from the early stage of the COVID-19 pandemic.

samples tested positive, particularly those from the sewage well and even the walls (Supplementary Table 2). Of the 70 RT-qPCR-positive samples, 36 samples (27 within the HSM and 9 from the surrounding area) collected in February were still positive for SARS-CoV-2. The long persistence of its genetic material in the environment might reflect high levels of environmental contamination before the market was closed. For sample Env_0838, collected from a wall on 20 February 2020, a 3-plex RT-qPCR test was carried out. The viral RNA segment was undetectable in one RT-qPCR channel targeting the *N* gene, but could be amplified in the other two channels targeting the *RdRp* and *E* genes, with Ct values of 32.59 and 37.34, respectively. This result is reasonable considering the degradation of the viral genome. However, the results also indicate a long persistence of the viral RNA in the environment.

We further carried out high-throughput sequencing (Supplementary Table 3) and successfully obtained seven complete or near-complete SARS-CoV-2 genome sequences, including three sequences from three environmental samples (Env_0313, Env_0354 and Env_0020), and four sequences from cell supernatants of Env_0313, Env_0354 and Env_0126 (Fig. 3 and Supplementary Table 4). A few samples were resequenced using a multiplex PCR approach, including Env_0020_seq01, Env_0313_seq04, Env_0313_seq05, Env_0126_seq06 and Env_0354_seq07 (Supplementary Tables 3 and 4). The genome sequences of three environmental samples, Env_0126, Env_0313 and Env_0354, were found to be identical to the reference strain HCoV-19/Wuhan/IVDC-HB-01/2019 (IVDC-HB-01, Global Initiative on Sharing All Influenza Data accession number: EPI_ISL_402119) and the human strain Wuhan-Hu-1 (GenBank: NC_045512; Fig. 3a). The genome sequence of the isolated virus from the environmental sample Env_0354 had two synonymous mutations compared to HCoV-19/Wuhan/IVDC-HB-01/2019, with sequence identity of 99.99% (Fig. 3a). Therefore, the SARS-CoV-2 sequences from environmental samples were highly similar to the clinical strains obtained during the early stages of the COVID-19 outbreak.

SARS-CoV-2 has been proposed to be classified into two main lineages based on the two highly linked single nucleotide polymorphisms: A lineage (8782T and 28144C, or S lineage in another nomenclature for SARS-CoV-2) and B lineage (8782C and 28144T, or L lineage). It has

been proposed that the A lineage is most probably the ancestral lineage, because all of the SARS-CoV-2-related coronaviruses from bats and pangolins possessed 8782T and 28144C (refs. 25,26); Pekar et al. suggested that the two lineages may represent separate introduction events[27]. Phylogenetic analysis revealed that most of the environmental strains belong to the B lineage and they cluster together with the human strains circulating in the early stage of the pandemic (Fig. 3b and Supplementary Fig. 1). The phylogenetic analysis did not involve the environmental sample Env_0020, the A lineage of which was confirmed by the high number of reads mapped to positions 8,782 and 28,144 in Env_0020 (Supplementary Table 5). However, it should be noted that the genome of Env_0020 is of low quality and there are many discontinuous gaps in the assembled genome. Indeed, although it is difficult to root the SARS-CoV-2 phylogenetic tree, our analysis indicated that the environmental viruses clustered together with the human strains circulating in the early stages of the pandemic.

We conducted RNA-sequencing (RNA-seq) analysis using 57 SARS-CoV-2 RT-qPCR-positive and 115 SARS-CoV-2 RT-qPCR-negative environmental samples from the HSM (Fig. 4a and Supplementary Table 3), in which the bias of sampling and RNA-seq should be considered. We used two approaches for identification of genera. The Kraken2 method with all available genes and genomes in the database was used for the identification of all genera, including those of Bacteria, viruses, Eukarya and Archaea. Additionally, the barcoding method using mitochondrial cytochrome *c* oxidase subunit sequences was used specifically for the identification of Chordata genera. Bacteria were the most abundant species in almost all samples and mammal species could be found in most samples, fitting the features of samples collected from the environment (Fig. 4b and Supplementary Tables 6 and 7). *Gallus*, *Homo*, *Anas*, *Sus*, *Bos* and *Canis* could be detected in most samples (Fig. 4c and Supplementary Table 8), in accordance with the environmental features of the seafood markets in China. We analysed the mammalian genera in all sequenced samples with Kraken2 (detailed in the Methods) using different thresholds. A total of 70 mammal genera, which existed in more than 2% of samples, were identified with a threshold of 100 reads per million (Fig. 4d). It is important to highlight that the results of the Kraken2 analysis (Fig. 4d) and the barcode of life data (BOLD) analysis (Extended Data Fig. 3) differ. In particular, the proportion of reads assigned as raccoon dog differs considerably between the two methods. This may be due to the heterogeneity of the reference data used by the two methods (mitochondria for BOLD; whole genome for Kraken2). It should be noted that the genera identified using current approaches might be updated with additional reference genomes. As such, this list is not definitive and further in-depth analysis with other methods will be required to provide more precise information regarding the wildlife species present at the market. In particular, it should be pointed out that our approach probably returned some false-positive assignments, particularly with the less-abundant genera (for example, *Ailuropoda*).

In particular, we analysed three samples (Env_0126, Env_0313 and Env_0354) collected on 1 January 2020 with high levels of SARS-CoV-2 (Ct value < 30; Fig. 4e). The identified mammal genera in the Env_0313 and Env_0354 samples were related to species in the general food market, such as *Homo*, *Ovis*, *Bos*, *Canis*, *Sus* and *Felis*. Many mammalian genera were observed in the Env_0126 sample, but the most abundant mammalian genera were also related to the general food market, including *Bos* (77.30%), *Ovis* (19.91%), *Homo* (0.77%) and *Bubalus* (0.57%). *Pipistrellus* (0.002%) and *Lutra* (0.001%) were also found in this sample, but at extremely low relative abundances, raising the possibility of false detection. Moreover, we also noted that only *Homo*, *Ovis*, *Bos* and *Sus* reads but not species related to wildlife were found in the Env_0020 sample, the one with A lineage.

We illustrated the top-ranked genera in four areas of the market where multiple SARS-CoV-2 RT-qPCR-positive samples were detected. As shown in Fig. 4f, the top-ranked genera in these areas were *Homo* or other genera that generally exist in food markets. We also noted that *Nyctereutes* could be found in shop 25 of street 8, and *Atelerix* and *Erinaceus* could be found in shops 15–17 of street 7 (Fig. 4f). These genera were detected in both SARS-CoV-2-positive and SARS-CoV-2-negative samples, and actually more often in negative ones (Supplementary Tables 6–9); thus, conclusions about whether these animals were infected with SARS-CoV-2 cannot be drawn.

We checked samples that might relate to wildlife, such as samples collected in the defeathering machine and areas with visible blood spots. The most abundant mammal genus of the defeathering machine sample (Env_0584) was *Canis* (Extended Data Fig 3). The most abundant mammal species of the visible blood spot sample (Env_0262) were *Bos*, *Sus*, *Ovis* and *Bison*, respectively (Extended Data Fig. 3). Additionally, we plotted the distribution of some genera of concern, including *Myotis*, *Erinaceus*, *Mustela*, *Nyctereutes*, *Rhizomys*, *Meles* and *Melogale*. Most of these samples were distributed in the West Zone of the market (Extended Data Fig. 4), where wildlife products were sold, but this also reflects the zone was much more intensively sampled and analysed by RNA-seq. The distributions of *Homo*, *Sus*, *Bos*, *Gallus* and *Anas* were also dominant in this area, which was near the areas enriched in SARS-CoV-2 RT-qPCR-positive samples. The repeated sampling of the locations with RT-qPCR-positive results may contribute some bias to the distribution analyses of areas enriched in SARS-CoV-2 RT-qPCR-positive samples. Additionally, we plotted the proportions of mammal genera in those SARS-CoV-2-positive samples with a high abundance of genera related to wildlife, such as Env_0576 (*Nyctereutes* enriched), Env_0807 (*Lariscus* enriched), Env_0809 (*Erinaceus* enriched) and Env_0585 (*Erinaceus* enriched; Extended Data Fig. 3).

Of particular note was the difference in the results from RT-qPCR and next-generation sequencing (NGS). As the RT-qPCR detection assay used in the very early stage of the pandemic was not formally verified, we believe that there may be some false positives and false negative in the RT-qPCR detection results in this study. We also found that SARS-CoV-2 reads could be detected by NGS in a portion of SARS-CoV-2 RT-qPCR-negative samples, possibly owing to degradation of SARS-CoV-2 within the RT-qPCR target region or contamination during library building. Additionally, we observed a relatively higher positivity rate when aligning the reads to the reference SARS-CoV-2 genome with Bowtie2 (Supplementary Table 6). Therefore, more precise algorithms are required to better capture the reads of SARS-CoV-2 RNA.

In summary, we report the detection of SARS-CoV-2 RNA and live virus in environmental samples from the West Zone of the HSM. It should be noted that the selection of shops for sampling was biased because shops selling wildlife as well as shops linked to early cases were prioritized for sampling. The origin of the virus cannot be determined from the analyses available so far. Although gene barcode analysis of animal species in the study suggested that *Myotis*, *Nyctereutes* and *Melogale*—species that have been recognized as potential host species of sarbecoviruses—were present at the market, these barcodes were mostly detected within the SARS-CoV-2 RT-qPCR-negative samples from the environment. It remains possible that the market may have acted as an amplifier of transmission owing to the high number of visitors every day, causing many of the initially identified infection clusters in the early stages of the outbreak[24].

Recent reports traced the outbreak back to the HSM and proposed, after compiling information reported by various sources, including the Joint WHO-China Study and social media, that the market sold live wild animals as recently as 2019 (ref. 28). Another report proposed that SARS-CoV-2 spilled over from animals to humans at least twice in November or December 2019, and the raccoon dog was suggested to be the intermediate host animal[27]. The evidence provided in this study is not sufficient to prove such a hypothesis. Our study confirmed the existence of raccoon dogs, and other potential SARS-CoV-2-susceptible animals, at the market before its closure. However, these environmental

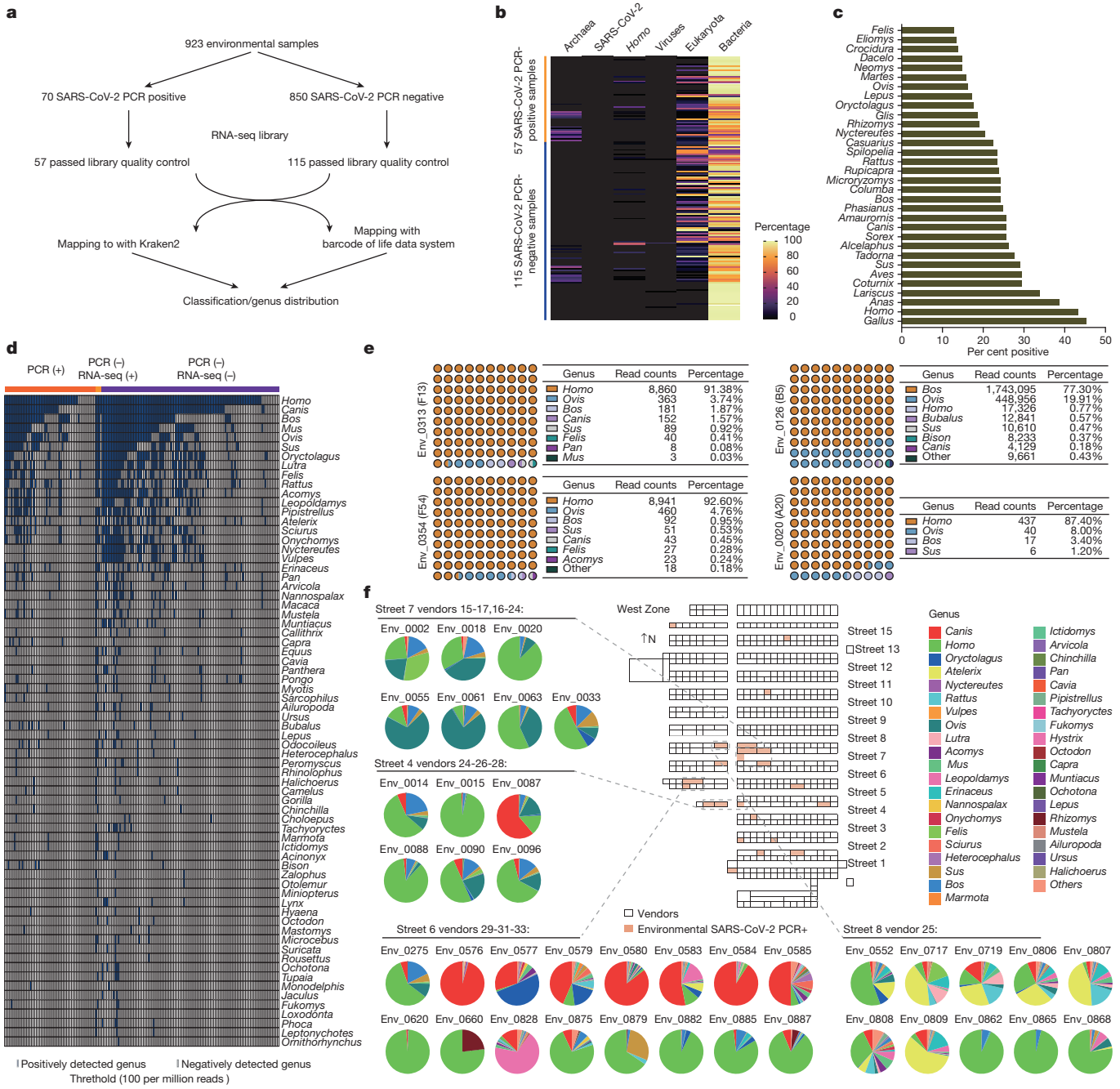


Fig. 4 | Analysis of environmental samples in the HSM. a, Schematic illustration of the experimental design. All 70 SARS-CoV-2-positive samples by RT-qPCR were included for RNA-seq. A total of 57 RNA-seq libraries were successfully constructed. Additionally, RNA-seq libraries of 115 SARS-CoV-2-negative samples passed library quality control. Kraken2 was used for genus classification. Kraken2 and the BOLD system were used for genus classification of Chordata. **b**, Heatmap showing the read distribution for four domains (Bacteria, Eukarya, viruses and Archaea), the *Homo* genus and the SARS-CoV-2 species, for SARS-CoV-2 RT-qPCR-positive or SARS-CoV-2 RT-qPCR-negative samples. **c**, Positive ratio of the illustrated genus in all tested samples. Top-ranked genera within the Chordata phylum are shown. **d**, Illustration of mammal genera in the market using the threshold of 100 reads per million based on Kraken2. The samples are grouped by SARS-CoV-2 RT-qPCR result and the NGS results analyzed with Bowie2. The blue bars indicate positively detected genera. **e**, Illustration of mammal genus distribution in samples with a high viral load. Data for Env_0020, Env_0313, Env_0354 and Env_0126 are shown. **f**, Distribution of the positively detected mammal genera in the market. Samples from four areas where multiple SARS-CoV-2 RT-qPCR-positive samples were detected are shown. The distribution of top mammal genera in each area is shown.

samples cannot prove that the animals were infected. Furthermore, even if the animals were infected, our study does not rule out human-to-animal transmission, as the sampling was carried out after the human infection within the market[6]. Thus, the possibility of potential introduction of the virus to the market through infected humans, or cold-chain products, cannot yet be ruled out.

More work, involving internationally coordinated efforts, is needed to investigate the potential origins of SARS-CoV-2 (ref. 24). Surveillance of wild animals should be enhanced to explore the potential natural and intermediate hosts for SARS-CoV-2 (refs. 7,29), if any, which would help to prevent future pandemics caused by coronaviruses of animal origin.

Note added in proof: The original, unedited Accelerated Article Preview (AAP) version of this Article contained some errors, which have been corrected in the final proof. In the AAP version, we provided the results of analysis performed in early 2020 and reported that SARS-CoV-2 was detected in 73 of the 923 environmental samples by RT–PCR. However, this was incorrect. SARS-CoV-2 had been detected in 70 of the samples by RT–PCR, and SARS-CoV-2 reads had been detected in an additional 3 samples (Env-0552, Env-0576 and Env-0585) by next-generation sequencing (NGS) followed by mapping the reads onto the reference genome NC_045512 using Bowtie2 (reads ≥ 1). At that time, in early 2020, only a few samples had been sequenced. We have now updated the Article to reflect the full set of sequencing results from all 172 samples that had sufficient RNA abundance for NGS and have indicated in the paper that 74 samples tested positive for SARS-CoV-2: 70 by RT–PCR and an additional 4 by NGS (using Bowtie2 analysis). Supplementary Table 1 has been modified to indicate which samples tested positive by RT–PCR or by Bowtie2 analysis. Supplementary Tables 2 and 5 have been revised to indicate that Env_0333 (F33) and Env_0509 (G93) did not pass quality control and NGS was not performed on these samples. We have also clarified in Supplementary Table 2 that samples Env_0552, Env_0576, Env_0585 and Env_0788 were negative by PCR but positive by sequencing and Bowtie2 analysis. Supplementary Table 6 has been revised to indicate which 33 samples were processed using the HWTSC002-16-BGI human rRNA depletion kit (BGI) and to include SARS-CoV-2 reads per sample. In the AAP version, we also incorrectly stated that sample Env_0354 (F54) was collected from a stall associated with a confirmed human case, instead of sample Env_0126 (B5). Consistent with Fig. 1a, samples Env_0313 (F13) and Env_0126 (B5), but not Env_0354 (F54), were from stalls with confirmed cases. We have also added further information to the Methods section to clarify details of the culture methods and sample processing for metagenomic sequencing.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-023-06043-2.

1. Tan, W. et al. A novel coronavirus genome identified in a cluster of pneumonia cases - Wuhan, China 2019-2020. *China CDC Wkly* **2**, 61–62 (2020).
2. Jiang, S. et al. A distinct name is needed for the new coronavirus. *Lancet* **395**, 949 (2020).
3. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **5**, 536–544 (2020).
4. Wang, C., Horby, P. W., Hayden, F. G. & Gao, G. F. A novel coronavirus outbreak of global health concern. *Lancet* **395**, 470–473 (2020).
5. Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
6. Li, Q. et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020).
7. Daszak, P., Olival, K. J. & Li, H. A strategy to prevent future epidemics similar to the 2019-nCoV outbreak. *Biosaf. Health* **2**, 6–8 (2020).
8. Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
9. Murakami, S. et al. Detection and characterization of bat sarbecovirus phylogenetically related to SARS-CoV-2, Japan. *Emerg. Infect. Dis.* **26**, 3025–3029 (2020).
10. Wacharapluesadee, S. et al. Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. *Nat. Commun.* **12**, 972 (2021).
11. Zhou, H. et al. A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein. *Curr. Biol.* **30**, 2196–2203 (2020).
12. Zhou, H. et al. Identification of novel bat coronaviruses sheds light on the evolutionary origins of SARS-CoV-2 and related viruses. *Cell* **184**, 4380–4391 (2021).
13. Li, J., Lai, S., Gao, G. F. & Shi, W. The emergence, genomic diversity and global spread of SARS-CoV-2. *Nature* **600**, 408–418 (2021).
14. Temmam, S. et al. Bat coronaviruses related to SARS-CoV-2 and infectious for human cells. *Nature* **604**, 330–336 (2022).
15. Lu, R. et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565–574 (2020).
16. Wang, J. et al. Individual bat viromes reveal the co-infection, spillover and emergence risk of potential zoonotic viruses. Preprint at *bioRxiv* https://doi.org/10.1101/2022.11.23.517609 (2022).
17. Lam, T. T. et al. Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature* **583**, 282–285 (2020).
18. Xiao, K. et al. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature* **583**, 286–289 (2020).
19. Niu, S. et al. Molecular basis of cross-species ACE2 interactions with SARS-CoV-2-like viruses of pangolin origin. *EMBO J.* **40**, e107786 (2021).
20. He, W. T. et al. Virome characterization of game animals in China reveals a spectrum of emerging pathogens. *Cell* **185**, 1117–1129 (2022).
21. Xiao, X., Newman, C., Buesching, C. D., Macdonald, D. W. & Zhou, Z. M. Animal sales from Wuhan wet markets immediately prior to the COVID-19 pandemic. *Sci. Rep.* **11**, 11898 (2021).
22. Wang, Q. et al. Tracing the origins of SARS-CoV-2: lessons learned from the past. *Cell Res.* **31**, 1139–1141 (2021).
23. Tong, Y. et al. The origins of viruses: discovery takes time, international resources, and cooperation. *Lancet* **398**, 1401–1402 (2021).
24. Joint WHO-China Study. *WHO-Convened Global Study of Origins of SARS-CoV-2: China Part* https://www.who.int/publications/i/item/who-convened-global-study-of-origins-of-sars-cov-2-china-part (WHO, 2021).
25. Tang, X. et al. On the origin and continuing evolution of SARS-CoV-2. *Natl Sci. Rev.* **7**, 1012–1023 (2020).
26. Rambaut, A. et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **5**, 1403–1407 (2020).
27. Pekar, J. E. et al. The molecular epidemiology of multiple zoonotic origins of SARS-CoV-2. *Science* **377**, 960–966 (2022).
28. Worobey, M. et al. The Huanan Seafood Wholesale Market in Wuhan was the early epicenter of the COVID-19 pandemic. *Science* **377**, 951–959 (2022).
29. Li, H. et al. Human-animal interactions and bat coronavirus spillover potential among rural residents in Southern China. *Biosaf. Health* **1**, 84–90 (2019).

Methods

Sample collection

The HSM was closed in the early morning of 1 January 2020, and at the same time, the China CDC began collecting environmental and animal samples. Staff from the China CDC entered the market about 30 times before the market's final clean-up on 2 March 2020, with some stray animals sampled outside the market until 30 March. Samples in the HSM were collected to represent as exhaustively as possible, from a wide diversity of surfaces, animals and products (Supplementary Tables 1 and 2 and Extended Data Table 6) according to different sampling principles, as described in detail in ref. 24.

The principles and ranges of in-market sampling covered: environmental samples from stalls related to early cases; environmental samples from doors and floors of all stalls in the blocks where the early cases were located; environmental samples, collected by block, from the East Zone of the market; transport carts, rubbish bins and similar objects; environmental samples from stalls that sold livestock, poultry or farmed wildlife (also referred to as domesticated wildlife or domesticated wildlife products); samples of sewage and silt from drainage channels and sewerage wells; stray cats, rats and other stray animals in the market; animal products and other commodity samples kept in cold storage and refrigerators in the market; the market's ventilation and air-conditioning system; and public toilets, public activity rooms and other places where people gathered in the market.

The investigators used full personal protective equipment during the sampling in the market. Commercially obtained swabs and virus preservation solution were used for the sampling (Disposable Virus Sampling Tube, V5-S-25, Shen Zhen Zi Jian Biotechnology). For environmental samples, sampling swabs were used to swab the floors, walls or surfaces of objects and then preserved in virus preservation solution.

For animal samples, depending on the type of animal and whether it was alive or frozen, pharyngeal, anal, body surface and body cavity swabs or tissue samples were collected for RT-qPCR. Generally, for live animals and frozen full bodies, three samples, including pharyngeal, anal and body surface swabs, were collected for each individual animal. For animal bodies after 'bai tiao' preparation (remaining parts of poultry or livestock after removal of hair and viscera), body cavity swabs were collected.

Drain samples were collected using virus sampling swabs to probe into the silt at the bottom of drainage channels in the market. Wastewater and silt samples were preserved in virus preservation solution. For the sewage well (for the drain water), a container was used to take a silt–water mixture from a location near the bottom of the well, and an appropriate amount of sample was collected by using virus sampling swabs and then preserved in virus preservation solution.

Nucleic acid extraction and SARS-CoV-2 RT-qPCR assay

A virus nucleic acid extraction kit (Xi'an Tianlong) was used to extract viral nucleic acid from samples using an automated nucleic acid extraction instrument according to the manufacturer's instructions. RT-qPCR was carried out on extracted nucleic acid samples with a SARS-CoV-2 nucleic acid assay kit. The reagent brands used include BioGerm (40/38; cycle number/cutoff value), DAAN (45/40) and BGI (40/38).

Virus isolation

Virus isolation was carried out in a biosafety level-3 laboratory in the National Institute for Viral Diseases Control and Prevention, China CDC. Samples positive for SARS-CoV-2 RT-qPCR collected on 1 January 2020 were cultured in both Vero E6 and Huh7.5 cells on 11 January 2020. The cells were cultured in 24-well cell culture plates with DMEM basal medium containing 10% fetal bovine serum and 1% penicillin-streptomycin in an incubator containing 5% CO_2. Homogenate supernatant was inoculated when the monolayer cell culture was about 90% confluent and adherent to the wall. The medium used was DMEM basal medium containing 2% fetal bovine serum. Three blind passages were carried out for each sample. The growth and morphological changes of the cells were observed under a microscope every day. The culture supernatant and cell pellet of each passage were collected for RT-qPCR. The morphology of viral particles in the cell sections and the supernatant were firstly observed by transmission electron microscopy, on 22 January 2020.

Metagenomic sequencing

Metagenomic sequencing was conducted at the National Institute for Viral Disease Control and Prevention, China CDC and Wuhan BGI. Nucleic acid was extracted using Qiagen's viral RNA microextraction kit. An enrichment kit (HWTS-C002-16-BGI, BGI, China) was used on 33 samples to improve the sensitivity of viral RNA detection. The kit is based on a probe pool that targets the human ribosomal RNA sequence. The probe pool comprises multiple oligonucleotide fragments, and viral RNA enrichment is accomplished through a sequence of steps including probe hybridization, RNAse H digestion, DNAse I digestion and magnetic bead purification. This specific treatment was chosen based on the low CT values (<30) of internal control (human genes) observed in these samples, indicating a relatively high abundance of human genes. However, the remaining samples did not undergo this treatment. Extracted RNA was reverse transcribed into cDNA and segmented into 150–200 base pairs by enzyme digestion. After repair, fitting, purification, PCR amplification and purification, the sample concentration was assayed by DNBSEQ-T7, and an average output of more than 200 million reads was obtained. Sequencing data were compared with those in a SARS-CoV-2 database to determine whether the samples contained SARS-CoV-2 sequences. For the seven complete SARS-CoV-2 genome sequences, three sequences from environmental samples (Env_0020_seq01, Env_0313_seq02 and Env_0354_seq03) were obtained from DNBSEQ-T7, and four sequences from cell supernatants of Env_0313, Env_0354 and Env_0126 (Fig. 3) were obtained from the NextSeq 550 platform. A few samples were resequenced using a multiplex PCR approach, including Env_0020_seq01, Env_0313_seq04, Env_0313_seq05, Env_0126_seq06 and Env_0354_seq07 (Supplementary Tables 3 and 4), as described previously[30]. Briefly, the nucleic acid was extracted using Qiagen's viral RNA microextraction kit. The multiplex PCR comprised a set of 102 oligonucleotide primer pairs and the amplicons generated by the primer pairs spanned the target genome. All raw data related to the genomes, including any partial genomes that were sequenced, were fully reported and deposited to the public database (Supplementary Tables 3 and 4).

Virus genome assembly and phylogenetic analysis

Raw reads were adaptor- and quality-trimmed with the Fastp (version 0.20.0) program. The clean reads were mapped to the SARS-CoV-2 reference genome (GenBank: NC_045512) using Bowtie2. The assembled genomes were merged and checked using Geneious (version 11.1.5) (https://www.geneious.com). The coverage and depth of genomes were calculated with SAMtools (version 1.10) based on SAM files from Bowtie2.

Reference genomes, IVDC-HB-01 (Global Initiative on Sharing All Influenza Data: EPI_ISL_402119) and Wuhan-Hu-1 (GenBank: NC_045512), were used as a query. Multiple sequence alignment of the SARS-CoV-2 sequences obtained from this study and reference sequences were carried out with Mafft (v7.450). Phylogenetic analyses were carried out using RAxML v8.2.9 with 1,000 bootstrap replicates, using the GTR nucleotide substitution model and the Gamma distribution.

Bioinformatic analysis of the species abundances

Kraken2 (version 2.1.2)[31] was used for species classification with the option --confidence 0.1. Sequences of all species in the Nucleotide (nt) database were used for generating the index. bracken (version

2.5) was used for re-evaluating species abundance. The matrix of species was obtained by using the pavian algorithm[32]. The ggplot2 package in R was used for plotting. Read counts of each genus were used for further analysis and plotting. Raw counts for four domains (Archaea, viruses, Eukarya and Bacteria), SARS-CoV-2 and the *Homo* genus were used to generate a heatmap (Fig. 4b). Two-tailed unpaired *t*-test was used for identification of differential genus between SARS-CoV-2 RT-qPCR-positive and SARS-CoV-2 RT-qPCR-negative samples.

For the analysis of the Chordata genus characterization, the reference was generated using the sequence of mitochondrial cytochrome *c* oxidase subunit I in the BOLD system[33–35]. RNA-seq samples were mapped to the reference sequences by the Bowtie2 (ref. 36) algorithm with the default settings. Read counts of each genus were calculated by samtools[37]. Read counts exceeding 20 were used as a cutoff for the identification of positively enriched genus. Fisher's exact test was used for comparing the differential genus in the Mammalia class between SARS-CoV-2 RT-qPCR-positive and SARS-CoV-2 RT-qPCR-negative samples.

Ethics

The sample collection was determined by the China CDC to be part of the emergency response to the outbreak of pneumonia of unknown aetiology and therefore was exempt from institutional review board assessment.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All of the raw sequencing data and genomes have been uploaded to the Global Initiative on Sharing All Influenza Data (China CDC Weekly, 2021, https://doi.org/10.46234/ccdcw2021.255). The accession codes are listed in Supplementary Tables 3 and 4. The raw sequence data reported in this paper have also been deposited in the Genome Sequence Archive (https://doi.org/10.1016/j.gpb.2021.08.001) in National Genomics Data Center (https://doi.org/10.1093/nar/gkab951), and the China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (publicly accessible at https://ngdc.cncb.ac.cn/gsa under the accession code CRA010170). The viral genomes reported in this paper have been deposited in GenBase at the National Genomics Data Center (Beijing Institute of Genomics, Chinese Academy of Sciences/China National Center for Bioinformation) under accession numbers C_AA002295.1 to C_AA002301.1, which are publicly accessible at https://ngdc.cncb.ac.cn/genbase/. Raw sequence data were also deposited into NCBI BioProject under accession number PRJNA948658 and in the China National Microbiology Data Center with the accession number NMDC10018366.

30. Niu, P. H. et al. Full-Length Genome Sequencing of SARS-CoV-2 Directly from Clinical and Environmental Samples Based on the Multiplex Polymerase Chain Reaction Method. *Biomed Environ Sci*. **34**, 725–728 (2021).
31. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol*. **20**, 257 (2019).
32. Breitwieser, F. P. & Salzberg, S. L. Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification. *Bioinformatics* **36**, 1303–1304 (2020).
33. Valentini, A., Pompanon, F. & Taberlet, P. DNA barcoding for ecologists. *Trends Ecol. Evol.* **24**, 110–117 (2009).
34. Hebert, P. D., Stoeckle, M. Y., Zemlak, T. S. & Francis, C. M. Identification of birds through DNA barcodes. *PLoS Biol*. **2**, e312 (2004).
35. Ratnasingham, S. & Hebert, P. D. bold: the Barcode of Life Data System (http://www.barcodinglife.org). *Mol. Ecol. Notes* **7**, 355–364 (2007).
36. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
37. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

Acknowledgements We acknowledge experts from Wuhan City, Hubei Province and across China who contributed to the study. We acknowledge the following experts for their contributions during this study: N. Mao and Y. Lan from the National Institute for Viral Disease Control and Prevention, China CDC; H. Jing and Q. Liu from the National Institute for Communicable Disease Control and Prevention, China CDC; L. Xu from Tsinghua University; Y. Jiang, J. Xu, X. Huo and B. Yu from Hubei Provincial CDC; Y. Xiong from Wuhan Municipal CDC; J. Li from Shandong First Medical University; W. Chen and H. Wu from BGI PathoGenesis Pharmaceutical Technology. In addition, we also acknowledge the work and suggestions of the joint team of scientists of the Joint WHO-China Study. W.J.L. is supported by the Excellent Young Scientist Program of the National Natural Science Foundation of China (81822040). The bioinformatics analyses by Y.T. in the study made use of the ASTRA computing platform in the National Research Center for Translational Medicine (Shanghai) and the Pi computing platform in the Center for High Performance Computing at Shanghai Jiao Tong University.

Author contributions The study was designed by G. Wu, W.J.L. and G.F.G. The onsite epidemiological survey and sample collection were carried out by W.J.L., W.L., Z.J., X.H., J.W., F.W., G. Wang, K.Q., R.G., Jie Zhang, M.L. W.X. and G.F.G. The nucleic acid extraction and RT-qPCR were carried out by W.J.L., P.L., W.L, Z.J., X.H., J.W., F.W. and G. Wang. NGS was carried out by W.J.L., P.L., W.L., Z.J., X.H., J.W., F.W., G. Wang and Wenting Zhou. Complete genome sequencing and analyses were carried out by P.L., Wenting Zhou, W.S. and W.J.L. The virus isolation was carried out by P.L., S.Z., Wenting Zhou, W.L., Jingdong Song and Z.X. Data analyses were carried out by W.J.L., P.L., Z.J., X.H., W.S., Y.T., S.Z., J.W., F.W., G. Wang, Y.G., Z.X., Y.Z., Juan Song, Jing Zhang, W. Zhen, Wenting Zhou, B.Y., Jingdong Song, M.Y., Weimin Zhou, Y.D., G.L., Y.B., W.T. and J.H. The manuscript was written by W.J.L., P.L., W.S., Y.T., G. Wong, G.F.G. and G. Wu.

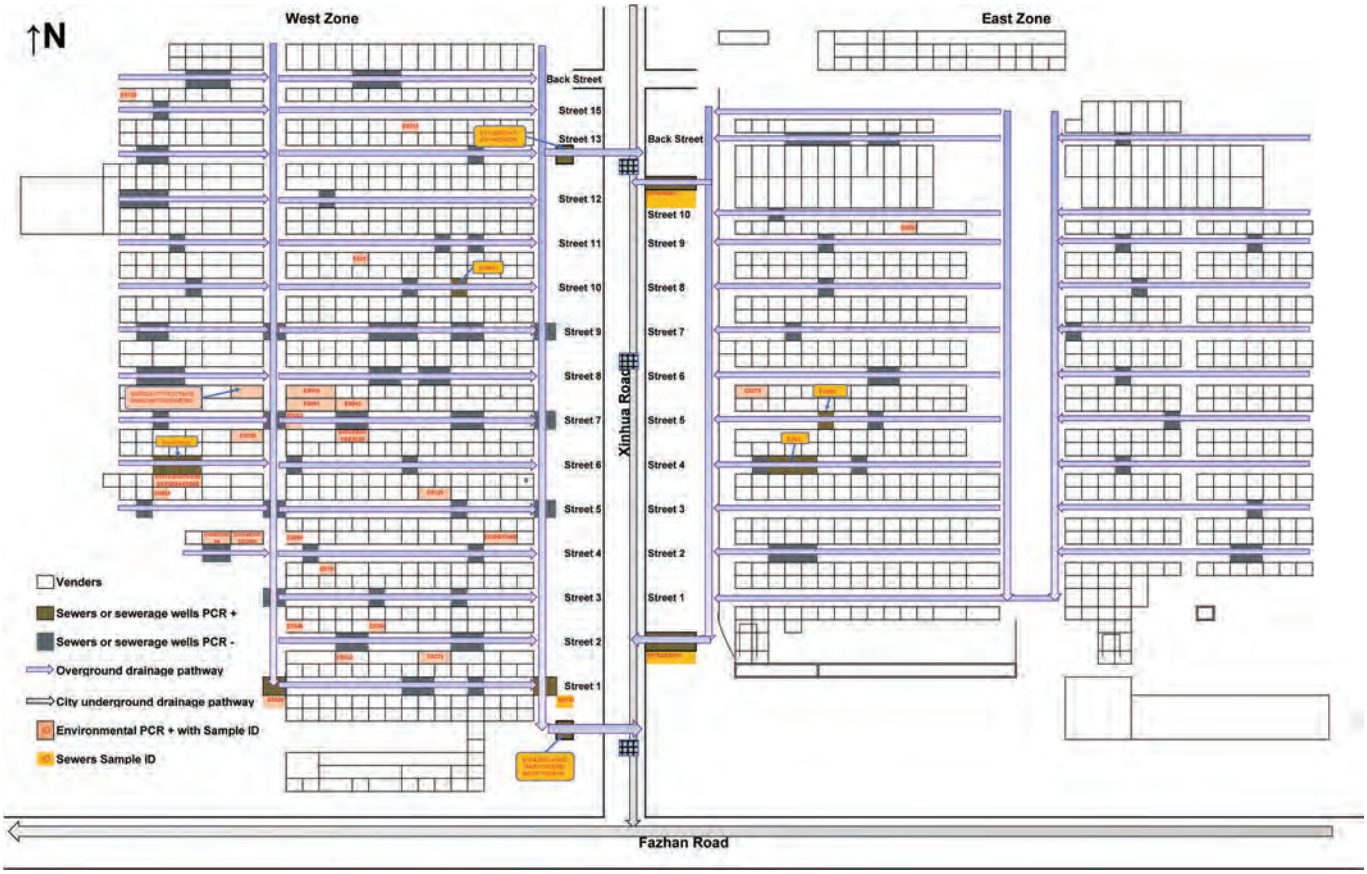
Competing interests The authors declare no competing interests.

Additional information

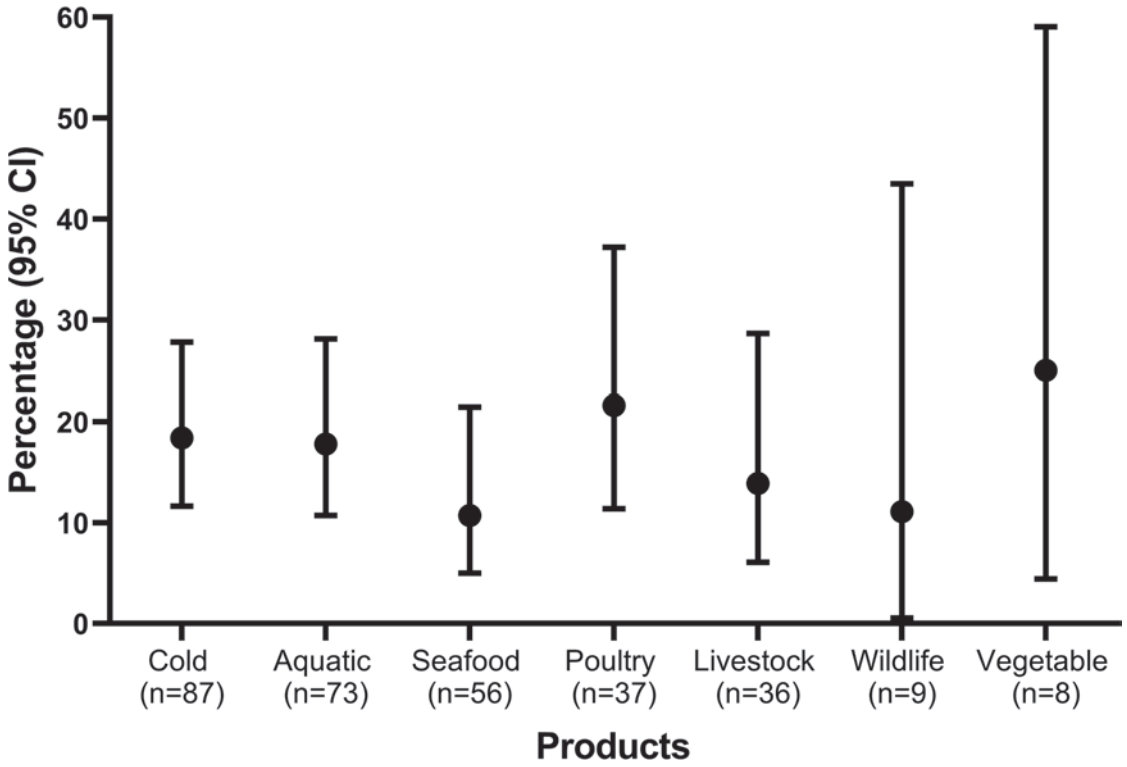
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-023-06043-2.

Correspondence and requests for materials should be addressed to William J. Liu, George F. Gao or Guizhen Wu.

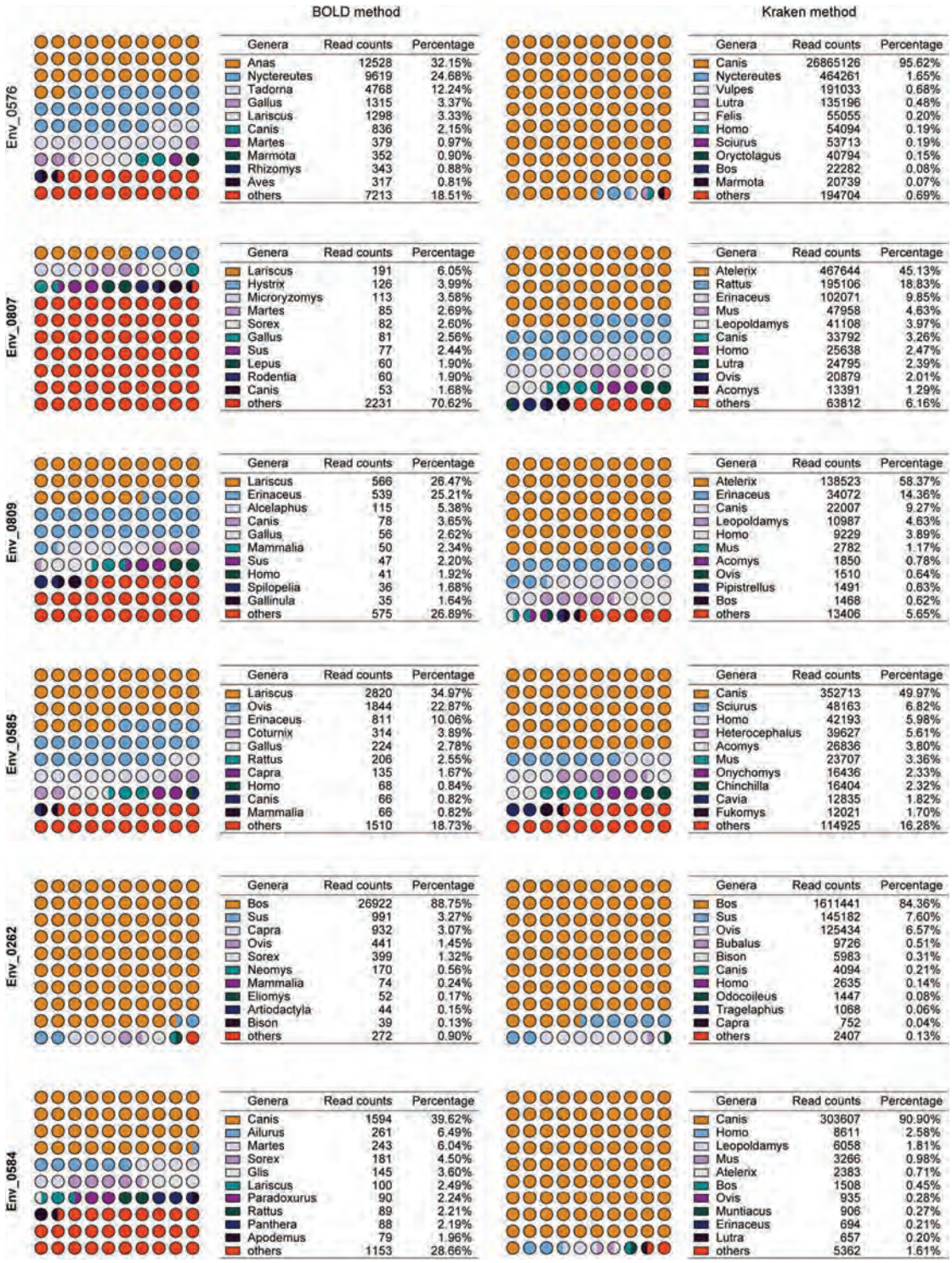
Peer review information *Nature* thanks Sébastien Calvignac-Spencer and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



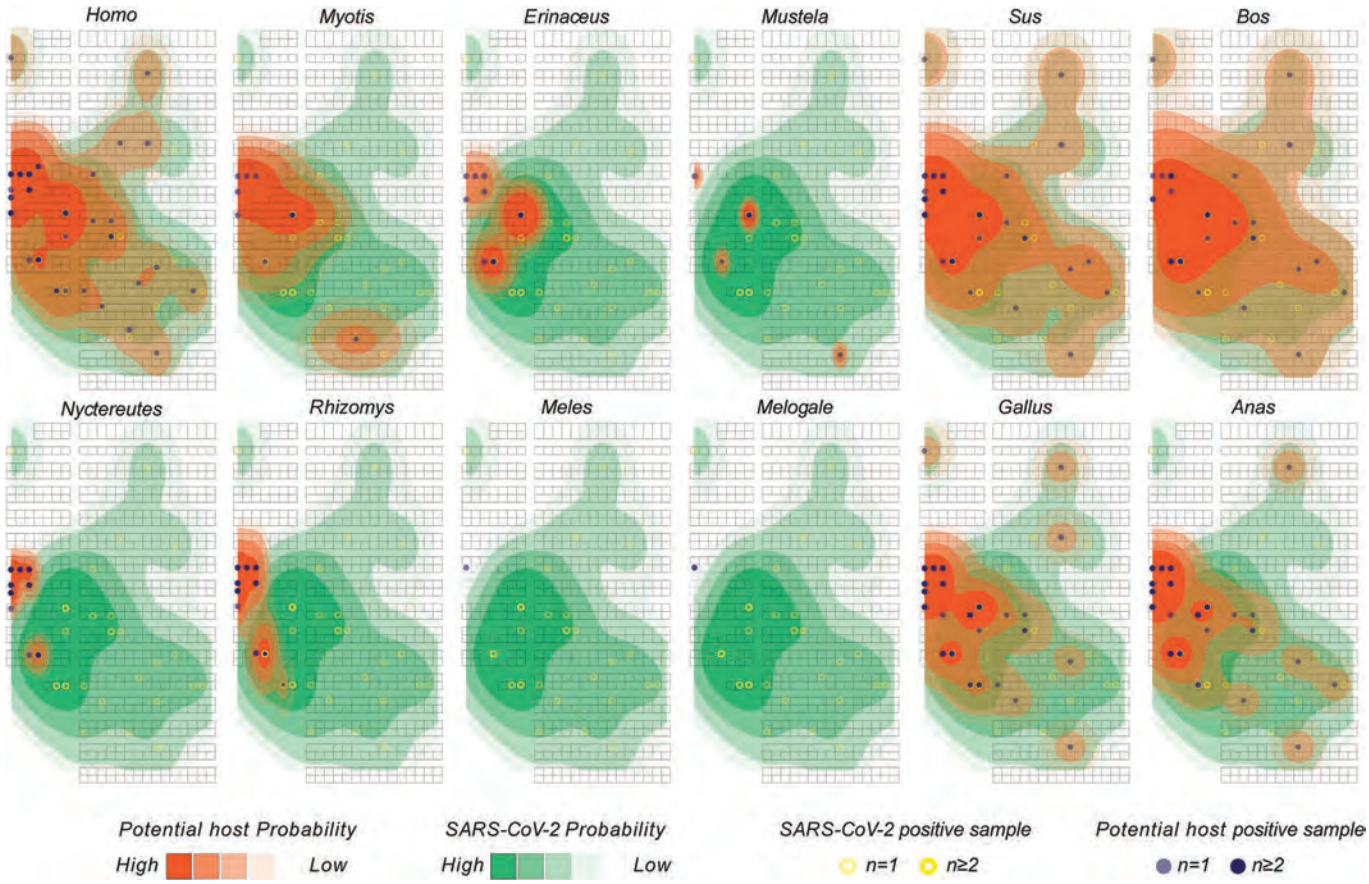
Extended Data Fig. 1 | The overground drainage pathway in the Huanan Seafood Market and environmental sample collection. The wastewater in the overground drainage was lead into the underground drainage inside the market and then flow into the wells on the edge of the market. And we did a spot-check sampling across all the overground drainages. To detect for the presence of SARS-CoV-2 RNA, RT-qPCR was performed. The locations of the positive samples were marked in the map of the market within yellow.



Extended Data Fig. 2 | Positive environmental samples associated with different products in the Huanan Seafood Market. Dots represent the percentage of positive environmental samples associated with each product. Bars represent 95% confidence intervals for the binomials in the text above. Note that the confidence interval (CI) for some products (e.g. vegetables, farmed wildlife) have broad error bars that are likely due to the low number of vendors for these categories in the market. Nine of the 10 vendors selling farmed wildlife have been sampled. Data are represented as percentage in this figure.



Extended Data Fig. 3 | Illustration of mammal genera distribution in samples of concerns. Illustration of mammal genera distribution in samples of concerns. Samples related to the blood spot and the de-feather machine (Env_0262 and Env_0584) and samples enriched with genera related to wildlife (Env_0576, Env_0807, Env_0809, and Env_0585) were plotted. Animal genera identified by the BOLD method were shown in the left panel, while mammal genera identified by the kraken2 method were shown in the right panel.



Extended Data Fig. 4 | Distribution of the positively detected Mammal genera in the market. The distribution of SARS-CoV-2 and potential host were plotted by yellow and blue dots, respectively. The density of the distribution of potential host was shown in red, while the SARS-CoV-2 by green.

Extended Data Table 1 | Overview of environmental sample sampling and testing in the Huanan Seafood Market

	Number of samples	Number of positive samples by RT-qPCR	Number of isolated viruses
Huanan Seafood Market	718	37	3
Warehouses related to the Huanan Seafood Market[a]	14	5	
Other markets in Wuhan and Huanggang[b]	30	1	
Drainage system in the Huanan Seafood Market	110	24	
Sewerage wells in surrounding areas	51	3	
Total	923	70	3

[a] The warehouses related to the Huanan Seafood Market were located out of the market.
[b] The one positive sample outside HSM was collected from Dongxihu Market in Wuhan. More information was provided in Extended Data Table 4.

Extended Data Table 2 | The collection logic of the environment samples

No.	Time	Objective	Sample time	Amount	Sum
1	1,Jan	(1) Environmental samples from stalls related to early cases; (2) Environmental samples from doors and floors of all stalls in the blocks where the early cases were located; (3) Environmental samples in the east wing of the market were collected according to blocks; (4) Transport carts, trash cans and similar objects.	1,Jan	515	515
2	12,Jan	Environmental samples from stalls that sold livestock, poultry, farmed wildlife (also called domesticated wildlife).	12,Jan	70	70
3	22,Jan	Environmental samples from other markets in Wuhan	22,Jan	30	30
4	23,Jan-19,Feb	The outdoor environmental samples from stalls that sold livestock, poultry, farmed wildlife.	23,Jan	23	52
			25,Jan	2	
			3,Feb	16	
			9,Feb	5	
			15,Feb	4	
			19,Feb	2	
5	27,Jan-15,Feb	Samples of sewage and silt from drainage channels and sewerage wells in the market.	27,Jan	38	94
			29,Jan	26	
			9,Feb	9	
			15,Feb	21	
6	5,Feb-9,Feb	Samples of sewage and silt from city sewerage wells around the market.	5,Feb	32	71
			9,Feb	39	
7	20,Feb-2,Mar	(1) Cold storages and refrigerators from stalls that sold livestock, poultry, farmed wildlife in the market; (2) The market's ventilation and air-conditioning system; (3) Public toilets, public activity rooms and other places where people gathered in the market.	20,Feb	27	91
			22,Feb	12	
			23,Feb	1	
			25,Feb	2	
			29,Feb	15	
			2,Mar	34	
Total				923	

Extended Data Table 3 | The collection logic of the animal samples

No.[a]	Time	Objectives	Sample time	Amount	Sum
8	22,Jan	Animal products in other markets.	22,Jan	6	6
9	25,Jan-10,Mar	Animal products and other commodity samples kept in the cold storages and refrigerators in the market.	25,Jan	55	306
			20,Feb	23	
			21,Feb	36	
			23,Feb	5	
			25,Feb	47	
			2,Mar	75	
			10,Mar	65	
10	27,Jan-1,Mar	Live animals captured around the market.	27,Jan	5	17
			5,Feb	3	
			9,Feb	2	
			15,Feb	3	
			29,Feb	2	
			1,Mar	2	
11	18,Jan-30,Mar	Stray cats, mice, cat feces and other stray animals (one dog and one weasel in the market).	18,Jan	1	96
			27,Jan	12	
			28,Jan	8	
			29,Jan	21	
			5,Feb	10	
			15,Feb	2	
			23,Feb	2	
			14,Mar	2	
			20,Mar	2	
			22,Mar	4	
			30,Mar	32	
12	19,Feb-23,Feb	Animal products and other commodity samples kept in the cold storages.	19,Feb	28	32
			23,Feb	4	
Total				457	

[a] The number follows the upper Table for environment samples.

Extended Data Table 4 | The information of the sampling in other markets

District	Number of environment samples [a]	Number of positive environment samples by RT-PCR	Number of animal samples [b]	Number of positive animal samples by RT-PCR
Jiang'han district	7	0	2	0
Jiang'an district	8	0	2	0
Donxihu district	7	1	1	0
Huanggang city	8	0	1	0
Total	30	1	6	0

[a] Swab sample collected from the floor, wall or chopping board.

[b] The heart, liver and large intestine tissues from pigs.

Extended Data Table 5 | Twenty-one shops of RT-qPCR positive in the Huanan Seafood Market

Vendors No.	Location	Product types[a]						
		Cold-chain products	Aquatic products	Seafood products	Poultry	Livestock	Wildlife products	Vegetables
1	West	no	no	no	yes	no	no	no
2	West	yes	yes	yes	no	no	no	no
3	West	yes	yes	no	yes	yes	yes	no
4	East	yes	no	no	yes	yes	no	no
5	West	no	no	no	no	no	no	no
6	West	no	yes	no	yes	yes	no	no
7	West	yes	no	no	yes	no	no	no
8	West	yes	yes	yes	yes	no	no	no
9	West	yes	yes	yes	no	no	no	no
10	West	yes	yes	yes	yes	yes	no	no
11	West	yes	yes	no	no	no	no	no
12	West	yes	yes	yes	no	no	no	no
13	West	yes	yes	no	no	no	no	no
14	West	yes	yes	no	no	no	no	no
15	West	yes	yes	no	no	no	no	no
16	West	yes	yes	no	no	no	no	no
17	West	no	no	no	no	no	no	no
18	West	yes	no	no	yes	yes	no	no
19	West	no	no	no	no	no	no	yes
20	West	yes	no	no	no	no	no	yes
21	East	yes	yes	yes	no	no	no	no
Sum of NAT positive vendors		16	13	6	8	5	1	2
Vendors sampled in the study selling such products		87	73	56	37	36	9	8

[a] "yes" indicates product sold by vendors; "no" indicates product not sold by vendors.

Extended Data Table 6 | The animal samples collected in the Huanan Seafood Market

Species	Animal number	Sample number	RT-PCR positive number
Rabbit/Hares	52	104	0
Stray cat	27	80[a]	0
Snake	40	80	0
Hedgehog	16	67	0
Muntjac	6	18	0
Dog	7[b]	17	0
Badger	6	16	0
Bamboo rat	6	15	0
Mouse	10	12	0
Pig	NA[c]	6[d]	0
Chicken	5	5	0
Chinese giant salamander	3	5	0
Crocodile	2	4	0
Wild boar	2	4	0
Soft-shelled turtle	2	3	0
Weasel[e]	1	2	0
Fish	2	2	0
Sheep	1	1	0
Others	NA[f]	16	0
Total	188	457	0

[a] Six of the cats were from the Huanan Seafood Market. And the samples included faeces.

[b] Including one stray dog in the Huanan Seafood Market.

[c] Not applicable due to the processed pork.

[d] Collected from other markets.

[e] The weasel was not sold in the market, but caught alive in the Market.

[f] Not applicable due to the unrecognized "bai tiao" product as described in the methods.